Review

Functional and structural insights into astacin metallopeptidases

F. Xavier Gomis-Rüth[1,*], Sergio Trillo-Muyo[1] and Walter Stöcker[2,*]

[1] Proteolysis Lab, Molecular Biology Institute of Barcelona, CSIC, Barcelona Science Park, Helix Building, c/Baldiri Reixac, 15-21, E-08028 Barcelona, Spain

[2] Institute of Zoology, Cell and Matrix Biology, Johannes Gutenberg University, Johannes-von-Müller-Weg 6, D-55128 Mainz, Germany

* Corresponding authors
  e-mail: fxgr@ibmb.csic.es; stoecker@uni-mainz.de

Abstract

The astacins are a family of multi-domain metallopeptidases with manifold functions in metabolism. They are either secreted or membrane-anchored and are regulated by being synthesized as inactive zymogens and also by co-localizing protein inhibitors. The distinct family members consist of N-terminal signal peptides and pro-segments, zinc-dependent catalytic domains, further downstream extracellular domains, transmembrane anchors, and cytosolic domains. The catalytic domains of four astacins and the zymogen of one of these have been structurally characterized and shown to comprise compact ~200-residue zinc-dependent moieties divided into an N-terminal and a C-terminal sub-domain by an active-site cleft. Astacins include an extended zinc-binding motif (HEXXHXXGXXH) which includes three metal ligands and groups them into the metzincin clan of metallopeptidases. In mature, unbound astacins, a conserved tyrosine acts as an additional zinc ligand, which is swung out upon substrate or inhibitor binding in a 'tyrosine switch' motion. Other characteristic structural elements of astacin catalytic domains are three large α-helices and a five-stranded β-sheet, as well as two or three disulfide bonds. The N-terminal pro-segments are variable in length and rather unstructured. They inhibit the catalytic zinc following an 'aspartate-switch' mechanism mediated by an aspartate embedded in a conserved motif (FXGD). Removal of the pro-segment uncovers a deep and extended active-site cleft, which in general shows preference for aspartate residues in the specificity pocket (S_1'). Furthermore, astacins undergo major rearrangement upon activation within an 'activation domain,' and show a slight hinge movement when binding substrates or inhibitors. In this review, we discuss the overall architecture of astacin catalytic domains and their involvement in function and zymogenic activation.

Keywords: bone morphogenetic protein; catalytic domain; meprin; metzincin; tolloid; zinc metallopeptidase.

Introduction: a short historical background

The first report on the digestive protease astacin from the European freshwater crayfish, *Astacus astacus* L. – then termed 'crayfish small-molecule protease' or '*Astacus* protease' – dates back to the late 1960s (Sonneborn et al., 1969). Protein sequencing by Zwilling and co-workers in the 1980s did not reveal homology to any other protein (Titani et al., 1987). Shortly after, the enzyme was identified as a zinc metallopeptidase (Stöcker et al., 1988), and other family members emerged. The first of these was bone morphogenetic protein 1 (BMP1), a protease co-purified with TGFβ-like growth factors termed bone morphogenetic proteins due to their capacity to induce ectopic bone formation in mice (Wozney et al., 1988). Later, prompted by the discovery of vertebrate meprins, the term 'astacins' was coined in 1991 to refer to a family of extracellular zinc endopeptidases encompassing them all (Dumermuth et al., 1991; Stöcker et al., 1991b). In the following years, cloning and sequence analysis of a variety of other astacins was reported, and most of them showed a multi-domain structure. They were shown to be involved in developmental processes, tissue differentiation, and embryonic hatching, as exemplified by UVS.2 from claw frog (Sato and Sargent, 1990), tolloid from fruit fly (Shimell et al., 1991), the low (LCE) and high (HCE) choriolytic enzymes from medaka fish (Yasumasu et al., 1992), and SPAN and blastula protein BP10 from sea urchin (Lepage et al., 1992; Reynolds et al., 1992).

Physiological background

In the human and mouse genomes, there are six genes encoding astacin proteases, namely, *bmp1*, *tll1*, *tll2*, *mep1a*, *mep1b*, and *astl* (see http://degradome.uniovi.es/met.html). The first three code for the tolloid subgroup, which includes protein BMP1 and its major splice variant, mammalian tolloid. These two are also known as procollagen C-proteases and are important for extracellular matrix assembly (Kessler et al., 1996; Li et al., 1996). Closely related to them are mammalian tolloid-like proteins 1 and 2 (TLL1 and TLL2), whose genes are differentially expressed when compared with that of BMP1. Knock-out mice for *bmp1*, *tll1*, and *tll2* have severe

defects in connective tissue assembly and heart and skeletal muscle development (for reviews, see Ge and Greenspan, 2006; Hopkins et al., 2007). These enzymes cleave precursors of fibrillar procollagens for proper matrix assembly. They also process other matrix proteins including proteoglycans, laminins, and anchoring fibrils. In addition, tolloids also cleave growth factors and their antagonists, which are crucial for dorso-ventral patterning during gastrulation in the embryo (Shimell et al., 1991; Holley et al., 1996; Ge and Greenspan, 2006).

Genes *mep1a* and *mep1b* encode the multi-domain proteins meprin α and meprin β, respectively. These are translated as membrane-bound proteins containing C-terminal MAM domains (meprin, A5 protein, and receptor protein tyrosine phosphatase μ) (Beckmann and Bork, 1993), TRAF domains (tumor necrosis factor receptor-associated factor) (Rothe et al., 1994; Zapata et al., 2001), and EGF-like, transmembrane, and cytosolic domains (Figure 1A). The α subunit is post-translationally cleaved within a unique I (inserted) domain (Figure 1A) during the passage through ER and Golgi, and therefore is found as high-molecular-weight soluble multimers. By contrast, meprin β homodimers and α/β heterodimers remain cell-surface-bound unless shed proteolytically (Hahn et al., 2003). Meprins are involved in tissue differentiation and pericellular signaling. In this context, a variety of meprin substrates have been reported *in vitro*, including biologically active peptides such as gastrin and cholecystokinin, substance P, cytokines, and chemokines (reviewed by Sterchi et al., 2008). Of special interest is the fact that meprins cleave components of the extracellular matrix, in particular the basal lamina but also adhesion proteins at the cell-cell interface (Sterchi et al., 2008; Ambort et al., 2010; Vazeille et al., 2011). Recent proteomics approaches have identified previously known and new physiologically relevant *in vivo* substrates such as vascular endothelial growth factor (Schütte et al., 2010), amyloid precursor protein (Jefferson et al., 2011), procollagens I and III (Kronenberg et al., 2010), interleukin-1β (Herzog et al., 2005), interleukin 18 (Banerjee and Bond, 2008), prokallikrein 7 (Ohler et al., 2010), and fibroblast growth factor 19 (Becker-Pauly et al., 2011).

The third subgroup of astacins in vertebrates comprises the so-called hatching enzymes, represented by just one member in mammals termed ovastacin by Carlos López-Otín and colleagues (Quesada et al., 2004). The protein is encoded by the gene *astl* and expressed in the oocyte and in the developing embryo. A recent report (Sachdev et al., 2012) suggested a role in sperm-egg interaction reminiscent of that of non-proteolytic members of the ADAM (a disintegrin and metalloproteinase) family of metalloproteases (Wolfsberg et al., 1993; Takeda, 2009; Takeda et al., 2012), which, like astacins, belong to the metzincin clan (Bode et al., 1993; Stöcker et al., 1995; Gomis-Rüth, 2003, 2009) (see chapter 'Overall structure of mature astacin catalytic domains' below).

Surprisingly, the genomes of lower vertebrates and invertebrates generally contain more astacin genes than mammalian genomes according to the MEROPS database (http://merops.sanger.ac.uk): 7–18 in amphibians and fish, at least 4 in cnidarians, 13–25 in insects, and up to 40 in nematodes such as *Caenorhabditis elegans* (Möhrlen et al., 2003). The large number in the latter example can at least partially be attributed to the parasitic lifestyle of nematodes, which requires an array of proteases to break down host connective tissue. Characterized examples are enzymes from *Trichinella spiralis* (Lun et al., 2003) and *Onchocerca volvolus* (Borchert et al., 2007). Also abundant in lower vertebrates and invertebrates are hatching enzymes, which degrade embryonic envelopes during the free water developmental stage of crustaceans, fish, frogs, and birds. The peptidases of this heterogeneous group often contain C-terminal cysteine-rich and CUB domains (complement C1r/C1s, Uegf, and BMP1) (Bork and Beckmann, 1993), yet some consist of just a single catalytic domain. Examples include crayfish embryonic astacin (Geier and Zwilling, 1998), fish alveolin (Shibata et al., 2000), fish proteins LCE and HCE (Yasumasu et al., 1996), nephrosin from carp head kidney (Hung et al., 1997), protein UVS.2 from frog (Fan and Katagiri, 2001), and protein CAM1 from bird (Elaroussi and DeLuca, 1994).

Modular organization of astacins and evolutionary aspects

The minimal structure of an astacin protease is a catalytic domain of approximately 200 amino acid residues as found in bacteria. In eukaryotes, this minimal structure is extended by at least an N-terminal pro-peptide which confers latency, so that most astacins, including the prototypical crayfish enzyme (Yiallouros et al., 2002), are secreted as inactive zymogens (Figure 1B). Additional downstream domains include EGF-like modules and/or one or more copies of CUB modules (Bond and Beynon, 1995; Stöcker and Bode, 1995). These domains may be involved in calcium-binding and protein-protein or enzyme-substrate interactions. It has been demonstrated that the C-terminal CUB- and EGF-like domains of procollagen C-peptidase are important for selectivity in substrate recognition (Sieron et al., 2000; Garrigue-Antar et al., 2004; Hintze et al., 2006; Wermter et al., 2007). By virtue of their similar, but not identical, domain composition, sea urchin astacins like SPAN and BP10 are related to tolloids (Lepage et al., 1992; Reynolds et al., 1992). They likewise contain a unique serine/threonine-rich region, which could be the target of O-glycosylation. Interestingly, some *C. elegans* astacins contain thrombospondin type 1 repeats, which are also found in another family of metzincins, the ADAMTS peptidases (Apte, 2009). Further C-terminal domains described include the ShK toxin domain of some coelenterate astacins. Such domains – also called six-cysteine (SXC) domains – were originally identified in metridin, a toxin from sea anemone, and several hypothetical *C. elegans* proteins. Other notable domains are the aforementioned TRAF and MAM domains (see section 'Physiological background'), found in meprins (see above) but also in HMP2 from hydra (Yan et al., 2000a,b) and LAST-MAM from the horseshoe crab (Becker-Pauly et al., 2009). Furthermore, regions of generally low compositional complexity and similarity to other protein modules
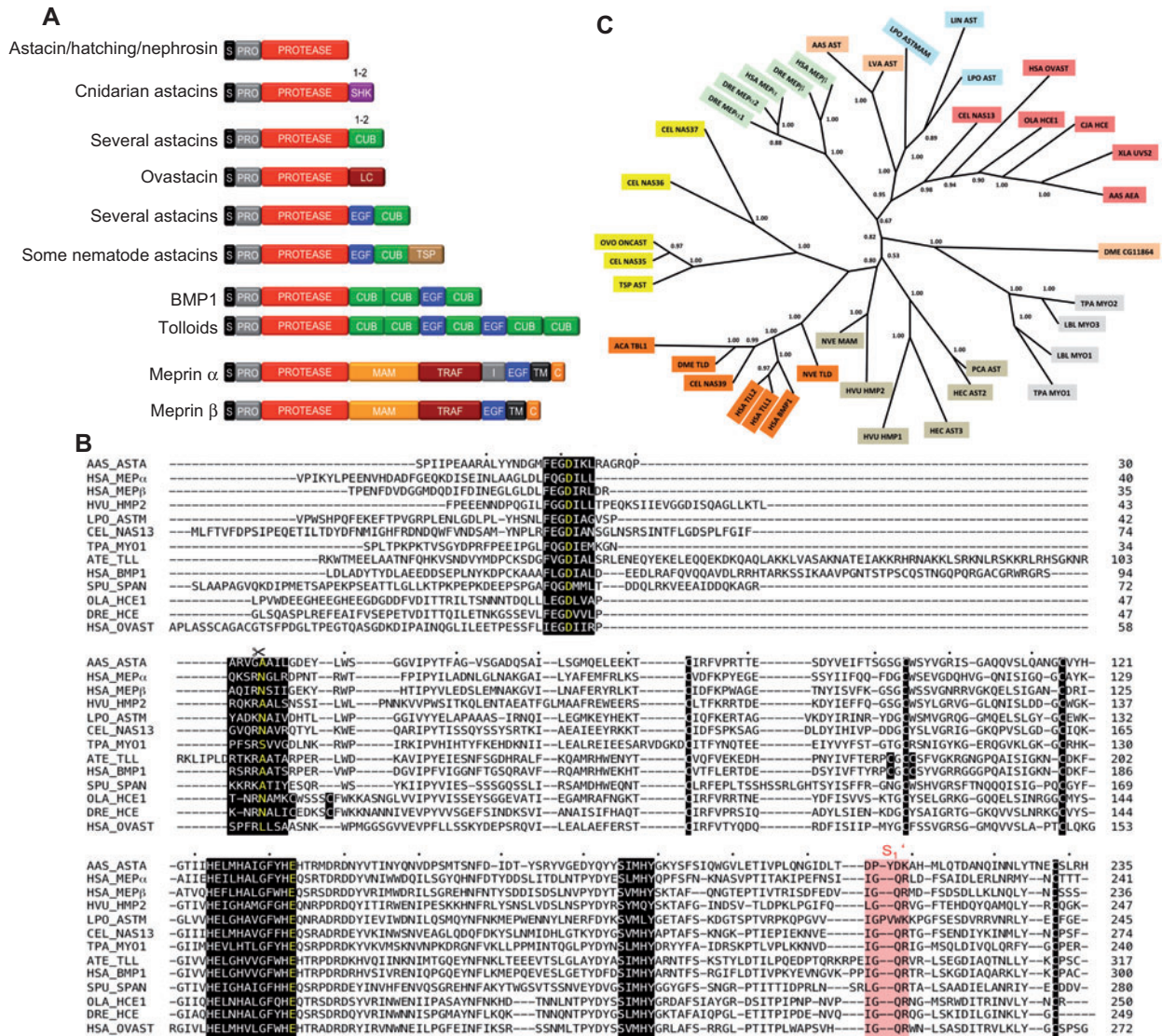
Figure 1 Architecture and evolution of astacins.

(A) Scheme depicting the distinct domains observed in astacins. In addition to N-terminal signal peptides and pro-peptides, most astacins contain further domains C-terminally attached to the catalytic protease domain. These are termed EGF (epidermal growth factor-like; PFAM accession number PF00008); CUB (named after their occurrence in complement component C1r/1s, embryonic sea urchin Uegf, and BMP1; PF00431); TSP (thrombospondin type 1 repeats; PF00090); ShK (ShK toxin domain; PF01549); EB (associated with Kunitz domains and found in several *C. elegans* proteins; PF01683); MAM (meprin, A5 receptor protein, tyrosine phosphatase μ; PF00629); TRAF (found in intracellular signaling proteins; PF00917); bacterial TT domains (PF02957); LC (low complexity domains); C (cytosolic domains); I (inserted domain); and TM (transmembrane anchor). (B) Sequence alignment of the pro-domains and catalytic domains of selected astacins. Over black background: the Met-turn, the zinc binding motif, cysteines, the aspartate-switch residue in the pro-peptide, and the activation site (scissors); in pink: residues chiefly shaping the S_1' sub-site. (C) Family tree based on the catalytic domains of astacins. The numbers indicate the relative probability of branching. UniProt database accession numbers: AAS AEA (O44072); AAS AST (P07584); ACA TBL1 (P91972); ATE_TLL (Q75UQ6); CEL NAS35 (P98060); CEL NAS36 (Q18206); CEL NAS37 (Q93243); CEL NAS39 (Q20176); CJA CAM1 (P42662); DME TLD (P25723); DME_AST (CG11864); DME_CG11864 (Q9VJN9); DRE_AST (=ZHE1) (Q75NR9); DRE MEPα1 (Q5RHM1); DRE MEPα2 (Q5RHM2); DRE MEPβ (Q08CC4); HEC AST2 (Q2MCX8); HEC AST3 (Q2MCX7); HSA BMP1 (P13497); HSA MEPα (Q16819); HSA MEPβ (Q16820); HSA OVAST (Q6HA08); HSA TLL1 (O43897); HSA TLL2 (Q9Y6L7); HVU HMP1 (Q25174); HVU HMP2 (Q9XZG0); LBL MYOI (Q8IU47); LBL MYOIII (Q8IU44); LIN AST (A0FKN6); LPO LAST (B4F319); LPO LAST_MAM (B4F320); NVE MAM (A7SJ13); NVE TLD (Q27W05); OLA HCE1 (P31580); OVO ONCHAST (Q2YFS7); PCA PMP1 (O62558); LVA AST (Q20AS7); SPU_SPAN (P98068); TPA MYOI (Q8IU46); TPA MYOII (Q8IU45); TSP MP (Q8T5Z5); and XLA UVS2 (P42664).

have been discovered and termed LC domains. Such regions have been observed in *C. elegans* astacins and sea urchin astacins SPAN and BP10 (Lepage et al., 1992; Reynolds et al., 1992). Moreover, mouse and human ovastacin contain a distinct C-terminal domain of approximately 150 residues with little similarity to other reported proteins (Figure 1A);

this domain might be heavily *O*-glycosylated (Quesada et al., 2004). Several other domains can be inferred from the more than 1000 astacin entries in the MEROPS database but they are not dealt with here as they have not been characterized at the protein level. Examples are the EB module found in some *C. elegans* proteins and the TT domain of bacterial astacins, which has been named after the viral ORF2 of the TT virus. The interested reader is referred to http://merops.sanger.ac.uk (Rawlings et al., 2010).

Alignment of representative pro- and catalytic domains of astacin peptidases reveals characteristic structural features which are associated with conserved functions (Figure 1B). There is the typical 'aspartate-switch' region in the pro-peptide (see chapter 'Zymogen structure and activation mechanism'), the zinc-binding consensus sequence, the 'Met-turn' (both discussed in the following section), and the unique S_1' sub-site, responsible for cleavage specificity (see chapter 'Active-site cleft, substrate specificity, and zinc-binding site'). A phylogenetic analysis based merely on the catalytic domains, and thus omitting pro-peptide regions and C-terminal domains, shows the interrelationships of astacin proteases (Figure 1C). There are several clusters of astacins. Some, like the tolloids, are present throughout the animal kingdom. Others, such as meprins, which have only been observed in vertebrates so far, are confined to distinct taxonomic groups. By contrast, hatching enzymes have diverged into many paralog lineages, especially in amphibians and fish. Other specialized subgroups of astacins seem to exist in nematodes, cnidarians, insects, and molluscs.

Overall structure of mature astacin catalytic domains

The first structure solved of a family member was that of crayfish astacin (Bode et al., 1992; Gomis-Rüth et al., 1993; Stöcker et al., 1993). It was the first metalloendopeptidase prototype to be structurally analyzed after thermolysin from *Bacillus thermoproteolyticus* (Matthews et al., 1972) and two closely related thermolysin-family members (Pauptit et al., 1988; Thayer et al., 1991). In contrast to other metallopeptidase (MP) groups such as matrix metalloproteinases (Tallant et al., 2010b), funnelin and cowrin metallocarboxypeptidases (Gomis-Rüth, 2008), the aforementioned thermolysins, and adamalysins/ADAMs (Takeda et al., 2012), few structures of astacins have been reported. To date, only human BMP1, human TLL1, and hatching enzyme 1 from the zebrafish *Danio rerio* (ZHE1) have been described in addition to the crayfish enzyme (Mac Sweeney et al., 2008; Okada et al., 2010). Furthermore, the structure of the astacin zymogen has also been published recently (Guevara et al., 2010). Overall, astacin catalytic domains (CDs) show a compact ellipsoidal shape, reminiscent of a kidney or a Pac-Man, with maximal dimensions of approximately 55×45×35 Å (Figure 2A). A deep and narrow active-site cleft divides the CDs into two sub-domains of approximately 100 residues when viewed in standard orientation (Gomis-Rüth et al., 2012), an upper N-terminal (NTS) and a lower C-terminal sub-domain (CTS;

Figure 2A–C). Superposition of BMP1, ZHE1, and TLL1 onto astacin reveals very similar chain traces, and this results in low overall *rmsd* values between these structures and astacin: 1.3 Å for BMP1 (for 174–180 equivalent Cα-atoms deviating <3 Å), 1.1 Å for TLL1 (175 equivalent Cα-atoms), and 1.0 Å for ZHE1 (178 equivalent Cα-atoms; Figure 2C). Accordingly, the main structural features, as well as residue numbering, will hereafter refer to astacin (see Bode et al., 1992; Gomis-Rüth et al., 1993) unless otherwise stated.

The NTS consists of a strongly twisted five-stranded β-sheet (β1–β5; connectivity -1×, +2×, +2, -1), whose strands parallel the active-site cleft except for its lowermost strand (β4) which creates the 'upper-rim' of the active-site cleft (see also the section 'Active-site cleft, substrate specificity, and zinc-binding site') and runs antiparallel (Figure 2B). The sheet is flanked on its top convex side by a long characteristic loop, which connects strands β$\underline{2}$ and β$\underline{3}$ (Lβ2β3), and by two helices on its bottom concave side, the 'backing helix' (αA) and the 'active-site helix' (αB), which run nearly parallel to the strands of the sheet. Helix αB includes the first three residues of the long zinc-binding consensus sequence $H^{92}EXXHXXGXXH^{102}$ (amino-acid one-letter code; X stands for any residue), which is characteristic of astacins but also metzincins in general (Bode et al., 1993; Stöcker et al., 1993, 1995; Gomis-Rüth, 2003, 2009). G^{99} within the consensus sequence is the endpoint of both helix α2 and NTS. This glycine allows for a sharp turn in the trajectory of the polypeptide in order to enter the CTS, and the values of its main-chain angles in the different structures ($\Phi=122°–140°$; $\Psi=10°–23°$) indicate that any other amino acid would be in a high-energy conformation (Davis et al., 2007). The CTS contains the third zinc-binding residue H^{102}, which is followed by the 'family-specific' residue of astacins (E^{103}) (Stöcker et al., 1993) (see also section 'Active-site cleft, substrate specificity, and zinc-binding site'). The rest of the CTS is characterized by few regular secondary structure elements, and only three short 3_{10}-helices (η1–η3) and two short strands (β6 and β7) are found in addition to the major 'C-terminal helix' (αC; Figure 2B). Special mention should be given to a tight 1,4-turn situated below the catalytic zinc-site, the Met-turn, which is characterized by a strictly conserved methionine (M^{147}), both in sequence and side-chain conformation, within astacins and also all other metzincins structurally analyzed to date (Gomis-Rüth, 2009; Goulas et al., 2010; Waltersperger et al., 2010). It has been proposed that the Met-turn acts as a plug that inserts laterally into a core structure created by the protein segment engaged in zinc binding, thus contributing to the structural integrity that is indispensable for function, but there is still debate on its significance in metzincins (Pieper et al., 1997; Boldt et al., 2001; Hege and Baumann, 2001; Butler et al., 2004; Walasek and Honek, 2005; Pérez et al., 2007; Oberholzer et al., 2009; Tallant et al., 2010a). Finally, a tyrosine two positions downstream of the methionine (Y^{149}) is also engaged in zinc binding and catalysis (see also the chapter 'Active-site cleft, substrate specificity, and zinc-binding site').
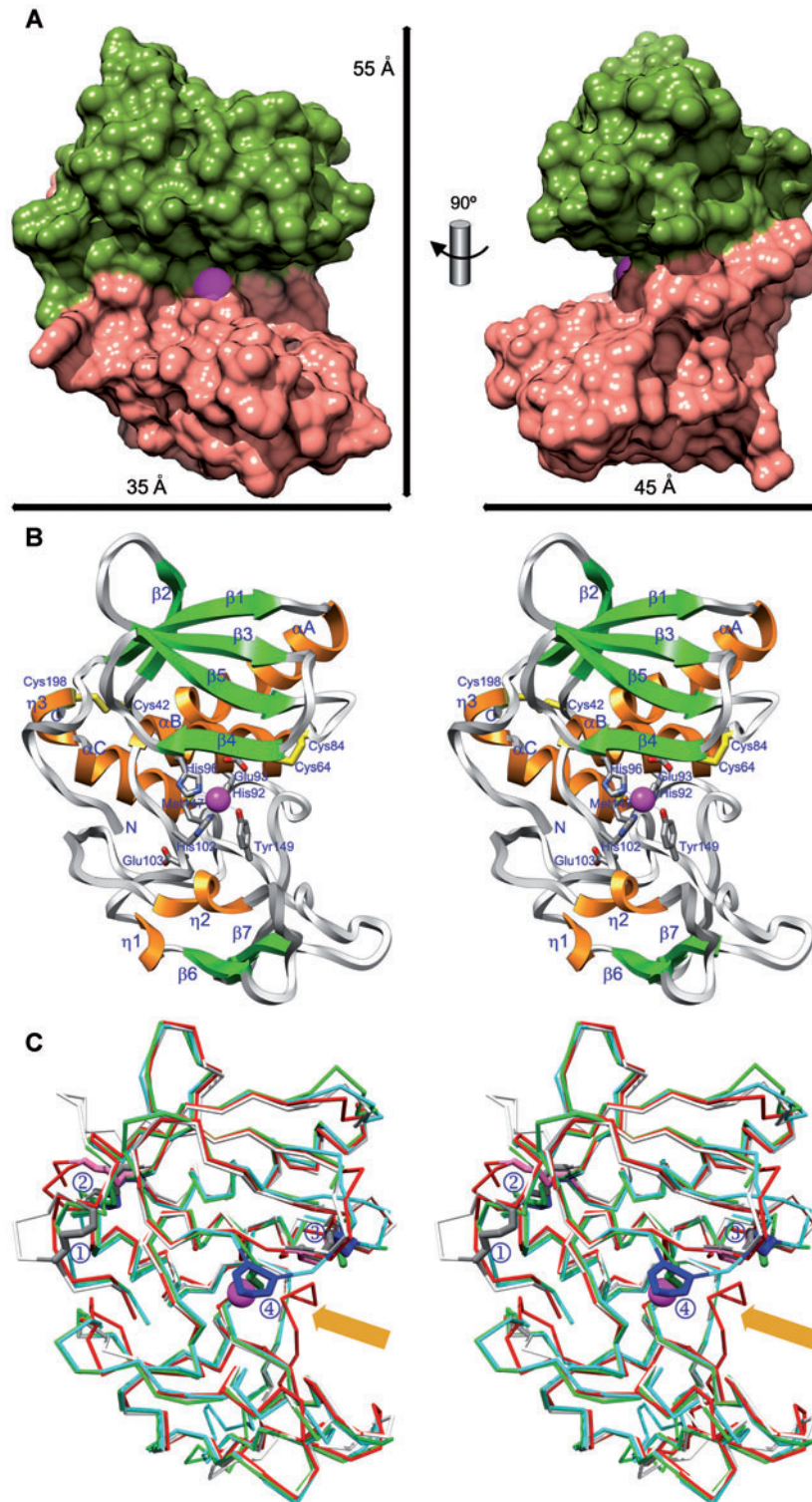
Figure 2   Overall structure of astacin catalytic domains.
(A) Mature astacin catalytic domain displayed with its Connolly surface in standard orientation according to Gomis-Rüth et al. (2012) (left) and after a vertical rotation of 90° (right). The NTS is shown in green, the CTS in salmon, and the zinc ion in magenta. (B) Richardson plot in stereo of astacin in standard orientation depicting its regular secondary structure elements as green arrows (β-strands β1–β7) and orange ribbons (α-helices αA–αC and 3_{10}-helices η1–η3). The two disulfide bonds are shown as yellow sticks and labeled, as are the catalytic zinc in magenta and its protein ligands, the termini, the Met-turn methionine, and the family-specific glutamate (E^{103}). (C) Superposition in stereo of the Cα-traces of astacin (red; pink disulfide bonds), BMP1 (cyan; dark blue disulfide bonds), TLL1 (green; dark green disulfide bonds), and ZHE1 (white; gray disulfide bonds). The four possible disulfide site locations are indicated by ①–④. The 170-loop is marked with an orange arrow.

Disulfide bonds

The CD of astacins is cross-linked by two or three disulfide bonds. Roughly common to all structures are the two found between C^{42} and C^{198} and between C^{64} and C^{84} in the crayfish prototype (Figure 1B and 2C; ② and ③). The former links the C-terminal segment of the CD to the NTS at the loop connecting helix αA with strand β2; the latter links Lβ3β4 and the beginning of β4 to Lβ5αB and thus contributes to shaping the active-site cleft at its primed site and to substrate binding (see also chapter 'Active-site cleft, substrate specificity, and zinc-binding site'). Sequence alignment with structurally non-characterized astacin family members (see Figure 1B and Stöcker et al., 1993) indicates that these two bonds are likely to be conserved among all astacins. In addition, selected members may show additional SS-bridges. ZHE1, as a member of the hatching subgroup, shows a unique cross-link between two cysteine residues in the N-terminal segment of the CD [C^5–C^{10} according to ZHE1 numbering; see Protein Data Bank (PDB) access code 3LQB; Okada et al., 2010; Figure 2C; ①]. This may be required to fix the very N-terminus in a competent position (see next section). BMP1 and TLL1, in turn, show a slight displacement of the disulfide bond at position ③ in Figure 2C and a further unique bond between two consecutive residues of the upper-rim strand β4 enabled by a *cis*-peptide bond between them (C^{64}–C^{65} according to BMP1 numbering; PDB 3EDH; Figure 2C; ④). This segment, termed 'cysteine-rich loop,' is unique for the tolloid subfamily (Figure 1C) and has implications for substrate binding (see section 'Active-site cleft, substrate specificity, and zinc-binding site').

A buried N-terminus in mature astacins

In general, proteins that undergo proteolytic maturation and possess additional flanking domains have their chain termini located on the molecular surface. While this is the case for the C-terminus of the CDs of mature astacins, the N-terminus is buried within the molecule for its first three or four residues (Figure 2B and C). In astacin, the first three residues A^1-A^2-I^3 are inserted like a plug in an internal cavity framed by segments D^{131}-Q^{142}, F^{100}-M^{107}, and T^{185}-Q^{189}. Residues from these segments, together with eight solvent molecules and the aforementioned N-terminal residues, are involved in an intricate, completely buried hydrogen-bonding network that is key for structural integrity of the enzyme (see Figure 3 of Bode et al., 1992) and incompatible with N-terminally elongated polypeptide chains. Most importantly, the α-amino group of A^1 establishes a solvent-mediated salt bridge with E^{103}, which is the 'family-specific' residue found immediately after the third zinc-binding residue (H^{102}; Figure 3). The solvent molecule further hydrogen-bonds the side chain of Q^{189}. These interactions contribute to structure and stability in astacin CDs. Mutants of pro-astacin, in which E^{103} had been replaced with glutamine and alanine, displayed unaltered catalytic efficiency but much lower thermal stability (Yiallouros et al., 2002). This structural rather than functional importance is supported by a superposable position and conformation of E^{103} in the zymogen and the mature structures (see also section 'Zymogen structure and activation mechanism').

By contrast, in both BMP1 and TLL1 structures, the first residue is an N-acetylated alanine. This means that the bridging solvent molecule is replaced by the carbonyl oxygen of the acetyl group, which is within hydrogen-binding distance of both Q^{189} and E^{103} (same numbering as in astacin; Mac Sweeney et al., 2008). In ZHE1, in turn, the structure of the mature enzyme starts at the position equivalent to A^2 of astacin, likewise with an alanine, so that its α-amino group is too far away from the E^{103}-equivalent to establish an interaction, and the empty space is occupied by six solvent molecules (Okada et al., 2010). However, the latter glutamate is maintained in a very similar side-chain conformation to that in astacin, TLL1, and BMP1, although in this case, an interaction takes places
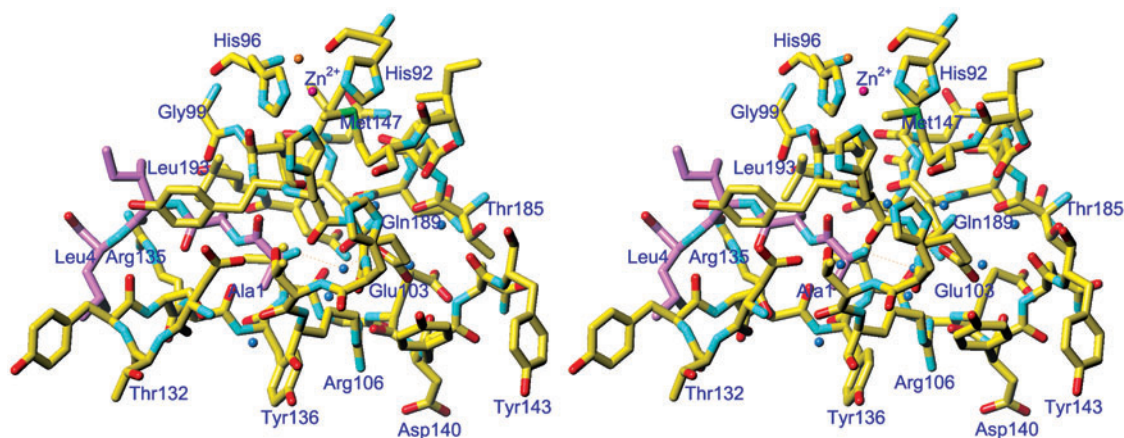
Figure 3 The mature N-terminus.
Stereographic picture centered on the buried N-terminus of the structure of astacin as a stick model colored according to atom types. For the N-terminal tail, carbons are shown in pink, otherwise in yellow. The zinc ion is depicted as a magenta sphere and labeled, and the eight solvent molecules of the internal cavity are shown as blue spheres. Selected residues are labeled for reference.

with an arginine at position 189, which replaces Q^{189} in the latter three proteins. This scenario also explains why ZHE1 possesses a unique disulfide bond that bridges the cysteine residues at positions 5 and 10 (see chapter 'Disulfide bonds' above), as this maintains the plugging function of the N-terminal segment despite the lack of the pivotal salt bridge with the family-specific residue. Inspection of the mature N-termini of representative astacin family members (Figure 1B and Figure 7 in Stöcker et al., 1993) shows that most structures should be compatible with the termini of either astacin/BMP1/TLL1 or ZHE1. N-terminal residues are almost exclusively alanine or asparagine, which is also compatible with the aforementioned structural features of a direct or water-mediated interaction with the family-specific residue. Overall, the presence of a buried N-terminus is unique to astacins within MPs and does not allow for variation in the length of the N-terminal segment which is essential for the maintenance of this specific structural feature. The latter finding is reminiscent of the internal salt bridge observed between the mature N-terminus of trypsin-like serine proteinases and an aspartate immediately adjacent to the catalytic serine residue (Fehlhammer et al., 1977; Huber and Bode, 1986).

Active-site cleft, substrate specificity, and zinc-binding site

In MPs, substrate hydrolysis proceeds via a high-energy reaction intermediate centered on a tetrahedral carbon bound to two *gem*-diolate oxygens, the scissile-bond nitrogen, and the preceding $C\alpha$ atom; the tetrahedral carbon itself is derived from the original scissile-bond carbonyl group by the nucleophilic attack of a solvent molecule polarized by a general base/acid glutamate (Matthews, 1988; Bayés et al., 2007; Gomis-Rüth, 2008). A complex of astacin with a reaction-intermediate analogue of peptide Phe-Pro-Lys-Phe-✀-Ala-Pro, in which the tetrahedral carbon and the downstream nitrogen flanking the scissile bond are replaced by a phosphinic group (Figure 4), allows us to delimit the deep and extended active-site cleft,

which horizontally traverses the catalytic domains of astacins. It explains why only elongated substrates – ideally longer that seven residues – are efficiently cleaved (Stöcker et al., 1990). In addition, comparison of this and other complexed forms of astacins with unbound enzymes further shows that astacins may undergo a slight overall hinge motion upon substrate, inhibitor, or ligand binding, which brings the CTS and the NTS closer by approximately 1 Å (Grams et al., 1996), in a fashion similar to that described for thermolysins (Holland et al., 1992).

In astacin and ZHE11, the upper-rim strand β4 frames the top of the cleft on its primed side, together with Lβ5αB and the disulfide bond at site ③ in Figure 2C. By contrast, in BMP1 and TLL1, this disulfide bond is slightly displaced and a further, unique SS-bridge is found between two consecutive cysteines within a cysteine-rich loop that replaces the upper-rim strand of astacin and ZHE1 (see section 'Disulfide bonds' above). This gives rise to an eight-membered, largely hydrophobic ring above the S_1 pocket (Figure 2C), which prevents substrate binding to the cleft and causes the upper rim to no longer be a β-strand. This cysteine-rich loop is disordered in the unbound structures, and it has been proposed to act as a mobile flap that clamps substrates into a competent position for a Michaelis complex (Mac Sweeney et al., 2008). On its non-primed side, the cleft is delimited in astacins by the end of strand β4 and the subsequent Lβ4β5, as well as the N-terminal segment at E^7-Y^8. At its bottom, the cleft is constrained on its non-primed side by I^4-G^5 and the loop after αB, and, on its primed side, by the Met-turn and the subsequent segment up to W^{158}, and, importantly, the '170 loop' mainly at T^{174}-D^{178}.

A substrate binds to astacins in an extended conformation and is anchored to the cleft in an antiparallel manner by the upper-rim strand β4 through inter-main-chain interactions, on both the primed and non-primed sides (Figure 4). The reaction-intermediate complex further reveals that cleft sub-site S_4 is framed in astacin by I^{71}, Y^8, I^3, V^{68}, and Y^{67}. In contrast, substrate residues in P_3 and P_1 protrude from the cleft toward the bulk solvent and lie up against the side chains of the upper-rim-strand residues W^{65} and Y^{67}. Sub-site S_2 is shaped by H^{96}, H^{102}, V^{68}, and Y^{101}. On the opposite side of the cleft, S_2' is created by
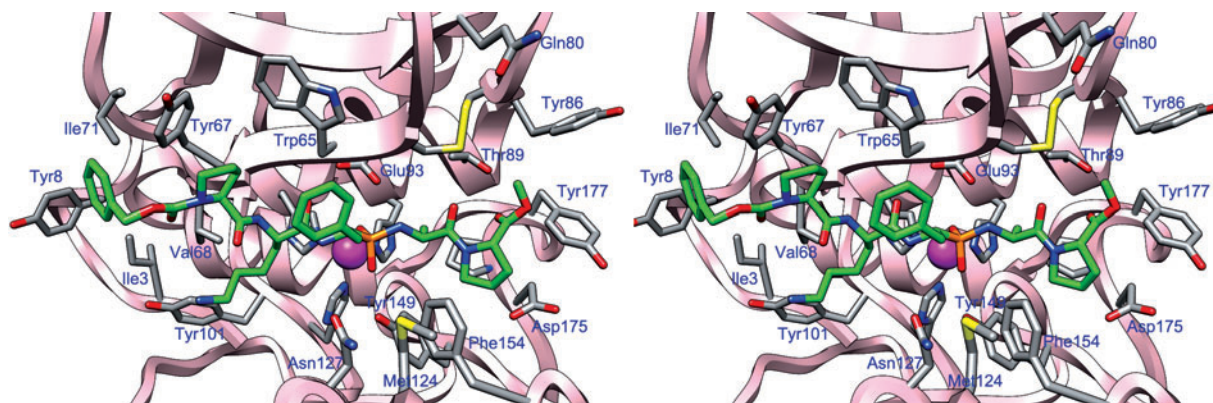
Figure 4 Substrate binding and enzymatic reaction.
Detail of the active-site cleft of astacin in stereo (pink ribbons with stick models of selected side chains colored according to atom types; carbons in gray) with the bound reaction-intermediate analog BOC-Pro-Lys-PheΨ(PO$_2$)-CH$_2$-C(CH$_3$)CO-Pro-OCH$_3$ (PDB 1QJI; Grams et al., 1996) likewise colored according to atom types (carbons in green, phosphorous in orange).

Y^{149}, D^{175}-P^{176}, and S^{153}-F^{154}; and additional downstream sub-sites are likely to be conditioned by D^{175}, Y^{177}, and W^{158}. As is usual for MPs, the most relevant sub-site for specificity is $S_1{}'$, which is delimited in astacin by H^{92} and the first helical turn of helix αB, in particular through the side chain of T^{89}. The most important structural element for primed sub-sites in astacins is the 170-loop, which shows disparate chain traces in ZHE1, TLL1, and BMP1, on the one hand, and astacin, on the other (see Figures 2C, 4 and Figure 7 in Okada et al., 2010). This leads to a conserved arginine (R^{182} in ZHE1 and R^{176} in BMP1/TLL1) shaping the bottom of a deep $S_1{}'$ pocket in ZHE1, TLL1, and BMP1, while the rather shallow pocket in astacin is shaped by P^{176}. This explains why most family members, including ZHE1, BMP1, TLL1, meprins α and β, and horseshoe-crab LAST-MAM, prefer aspartate residues in $P_1{}'$ while astacin prefers small aliphatic residues (Stöcker et al., 1990; Okada et al., 2010; Becker-Pauly et al., 2011). Accordingly, astacins, with the notable exception of their founding prototype and a minority of other family members, are the only eukaryote proteases to cleave N-terminal to an aspartate in proteins (Becker-Pauly et al., 2011). This preference is seen even more strongly in enzymes like meprin β and tolloids (e.g., BMP1), in which the $S_2{}'$ sub-site (see above) also contributes to the binding of acidic substrate side chains with basic residues.

The catalytic zinc ion resides at the bottom of the active-site cleft (Figure 2A and B). The active-site helix αB provides two histidine ligands of the metal (H^{92} and H^{96}), separated by a single helical turn that allows a concerted approach to the metal together with the third zinc-binding histidine, H^{102}. In unbound astacin, the Nϵ2 atoms of these residues coordinate the metal together with a catalytic solvent molecule, also anchored to the general base/acid E^{93}, and Y^{149} Oη (Gomis-Rüth et al., 1993). The overall metal coordination is trigonal-bipyramidal, with H^{96} Nϵ2 and Y^{149} Oη at apical, somewhat larger distances (2.2–2.5 Å), and the remaining three ligands on a plane with the metal, ~2 Å apart (Figure 12 in Gomis-Rüth et al., 1993). In ligand- or inhibitor-bound astacins, the catalytic solvent is replaced by one or two oxygens, and the Y^{149} ligand is pulled away from its metal-binding position. In the complex with the reaction-intermediate analogue, the Oη atom stabilizes one of the *gem*-diolate oxygen atoms of the tetrahedral carbon (Figure 4). This motion has been termed a 'tyrosine switch' and is also found in serralysins and – most likely – pappalysin family members within the metzincins (Gomis-Rüth, 2003, 2009; Baumann, 2004; Tallant et al., 2006). Finally, studies on apo-astacin and metal-replaced enzyme revealed that the metal site is preformed and that it can accommodate the ligand requirements of distinct divalent metals through minor rearrangement and recruitment of a variable number of solvent molecules. Overall, this gives rise to bipyramidal-trigonal, tetrahedral, or octahedral coordination spheres (Gomis-Rüth et al., 1994).

Zymogen structure and activation mechanism

Virtually all eukaryotic astacins are synthesized as inactive precursors, which prevents the occurrence of temporally and spatially inappropriate proteolytic activity. As an example, pro-astacin is only transiently found within the ducts between the hepatopancreas – where it is synthesized – and the stomach, where it is activated in order to participate in collagenolysis and gelatinolysis during digestion (Yiallouros et al., 2002). Comparison of the pro-segments of distinct family members (see Figure 1B and Figure 1 in Guevara et al., 2010) revealed that these vary greatly in length (from 34 to 393 residues) and that only a short consensus sequence $F^{18P}XGD^{21P}$ (residues of the astacin pro-segment carry the suffix 'P') is revealed by sequence alignments.

The only structure of an astacin-family zymogen reported to date is that of the crayfish enzyme (Guevara et al., 2010; PDB 3LQ0; Figure 5A). With merely 34 residues, pro-astacin possesses the shortest pro-segment structurally analyzed for an MP and, in contrast with other peptidases, it is not required for proper folding in the crayfish enzyme, which could be purified from *Escherichia coli* inclusion bodies and correctly folded as a recombinant mature protease, devoid of the pro-peptide (Reyda et al., 1999; Yiallouros et al., 2000, 2002). Of course, this might be different in pro-astacins with longer pro-domains, which in some cases like *Drosophila* tolloid-like are much longer than the catalytic protease domain itself. In pro-astacin, the pro-segment is elongated and structured by means of several intramolecular interactions, and it runs across the front surface of the mature enzyme moiety in the opposite orientation to that of a substrate (Figure 5A–C). This contributes to the prevention of self-cleavage, as found in cysteine-protease and matrix metalloproteinase zymogens (Khan and James, 1998). The N-terminus of the pro-segment at S^{1P} is anchored to the mature part through a hydrophilic interaction with the main chain of I^{156}, within the segment connecting the Met-turn and the C-terminal helix αC. The polypeptide runs along the molecular surface toward the active-site cleft and adopts a helical conformation from E^{6P} to Y^{12P} (helix $\alpha 1$ in Figure 5A). This segment nestles in the primed side of the cleft approximately until sub-site $S_1{}'$, which is partially occupied by the side chain of L^{11P}, and this contributes, together with a flexible segment within the CTS of the mature moiety of the zymogen (the 'activation-domain'; D^{129}-G^{138}; see below), to a larger separation between the two sub-domains than in the unbound mature enzyme. At N^{14P}, the chain projects toward bulk solvent and enters a wide loop that ends at D^{21P} (Figure 5A). This loop comprises two successive 1,4-turns, which together with hydrophobic interactions of M^{17P} with both K^{23P} and W^{65} give rise to a small globular nucleus, which places D^{21P} in contact with the catalytic zinc ion (Figure 5B and C). The loop structure is further stabilized by electrostatic interactions of E^{19P} with the zinc-binding residue H^{102}, T^{105}, and, most relevantly, the penultimate residue of the pro-segment (R^{32P}; Figure 5C). D^{21P} approaches the catalytic metal from the top in a bidentate manner and its Oδ2 atom substitutes for the zinc-bound solvent molecule in the unbound mature enzyme. Overall, the zinc ion shows a distorted octahedral coordination sphere (Figure 5B), which is unusual for zinc (McCall et al., 2000) and is reminiscent of the structure of the catalytically inert nickel-substituted
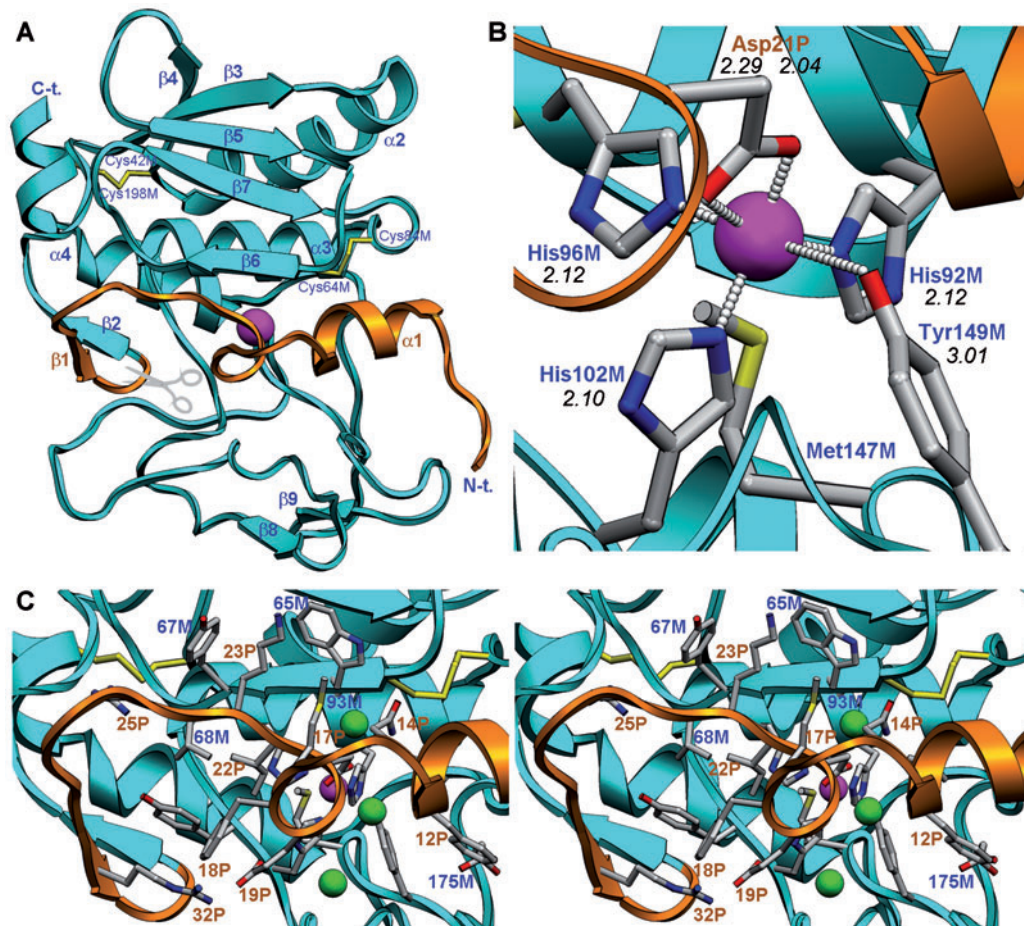
Figure 5 Structural determinants of zymogenicity.
(A) Richardson-type plot of pro-astacin in standard orientation (Gomis-Rüth et al., 2012) depicting the pro-segment in orange and the mature enzyme moiety in cyan. Repetitive secondary structure elements are shown as ribbons (helices α1–α4) and arrows (strands β1–β9), the catalytic zinc ion as a magenta sphere, and the two disulfide bonds as yellow sticks. The latter are labeled, as are the termini of the molecule. The final activation cleavage site (G^{34P}-A^{1M}; mature enzyme residues within the zymogen structure carry suffix 'M,' pro-segment residues carry suffix 'P') is shown by the scissors. For clarity, only one conformation has been displayed for segment D^{129M}-P^{135M}. (B) Close-up view of (A) depicting the catalytic zinc ion with its six ligands, which are labeled. The respective bonding distances (in Å) are shown below each residue label. The Met-turn methionine is also shown and labeled. (C) Close-up view of (A) in stereo to illustrate the major interactions between the pro-segment and the mature enzyme moiety. Participating residues are labeled (mature enzyme residues in blue, pro-segment residues in brown), except those already labeled in (B). Relevant solvent molecules are displayed as green spheres (reproduced from Guevara et al., 2010 © The American Society for Biochemistry and Molecular Biology).

astacin (Gomis-Rüth et al., 1994). By contrast, the side chain of tyrosine-switch residue Y^{149} is closer to the competent conformation of the unbound mature structure than to the structure bound to the reaction-intermediate analog, although somewhat further from the metal ion (see Figure 5B and Figure 12 in Gomis-Rüth et al., 1993). The four residues after D^{21P} run in extended conformation along the cleft between sub-sites S_2 and S_5 and bind the upper-rim strand (strand β6 in the pro-astacin structure; see Figure 5A and C). The polypeptide chain reaches the molecular surface after R^{25P}, whose side chain is anchored to the mature enzyme backbone, and thereafter, a flexible segment leads to a short β-hairpin structure made up by β1 (A^{31P}-V^{33P}) and β2 (A^2-L^4) (see Figure 5A and C). The main maturation cleavage point (G^{34P}-✂-A^1), which is buried within the molecular structure

of the zymogen like the N-terminus of the mature form (see section 'A buried N-terminus in mature astacins'), is found at the tip of this hairpin.

Activation of pro-astacin entails removal of the inhibiting pro-segment through successive cleavage events, which eventually replace the zinc-binding aspartate with the catalytic solvent molecule following an aspartate-switch mechanism and render the mature N-terminus at A^1 (Yiallouros et al., 2002; Guevara et al., 2010). In the first activation step, successive exogenous cleavages render an intermediate species starting at V^{33} (according to Yiallouros et al., 2002), which disrupts the aforementioned salt bridge (R^{32P}-E^{19P}) (see above), and triggers removal of most of the pro-segment. This causes major rearrangement of the subjacent 'activation domain,' which adopts the rigid and competent conformation
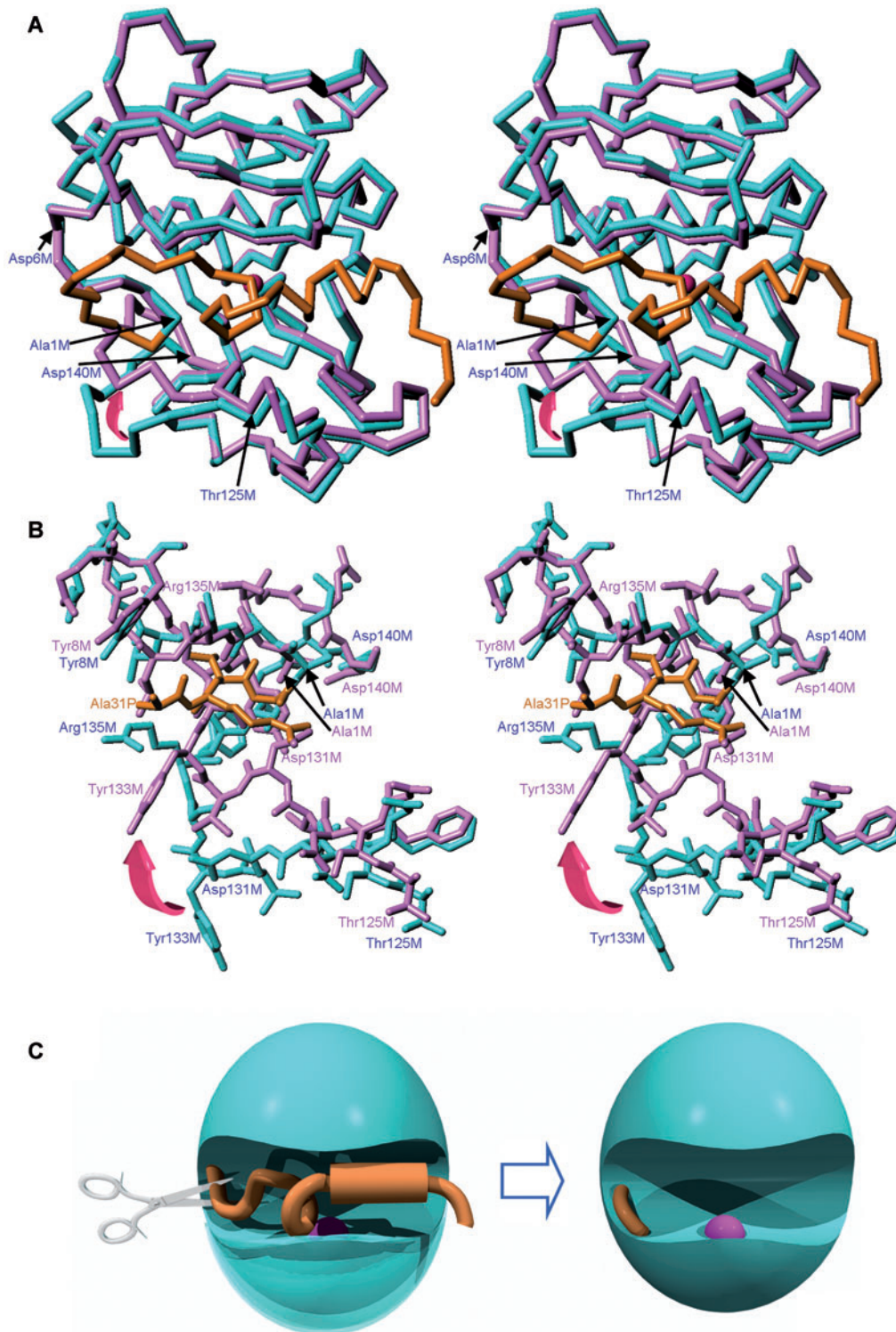
Figure 6 Structural transitions upon activation.

(A) Superimposition in stereo of the Cα-traces of pro-astacin (pro-segment in orange, catalytic moiety in cyan) and mature astacin (purple) in standard orientation (Gomis-Rüth et al., 2012). The catalytic zinc ion of pro-astacin is shown as a magenta sphere for reference. Segments involved in activation (magenta arrow) are indicated. Only one conformation is shown for segment D^{129M}-P^{135M} (mature enzyme residues within the zymogen structure carry suffix 'M'; pro-segment residues carry suffix 'P'). (B) Close-up view of (A) in stereo showing only the first residues of the mature moiety and the activation domain of both the pro-enzyme (cyan sticks; blue labels) and the mature enzyme (purple sticks and labels), as well as the four last pro-enzyme residues (orange sticks and labels). (C) Schematic representation illustrating the transition between the zymogen, with a flexible activation domain in the lower sub-domain of the molecule, and the rigid mature enzyme (reproduced from Guevara et al., 2010 © The American Society for Biochemistry and Molecular Biology).

found in the mature enzyme (Figure 6A–C). This is analogous to serine proteinases, in which the activation domain likewise becomes rigid and functional upon activation (Bode and Huber, 1978; Khan and James, 1998). In addition, the initial multiple exogenous cleavages are reminiscent of matrix metalloproteinase activation during which trimming cuts by other proteinases in a so-called 'bait region' are observed (Nagase, 1997). This contrasts with funnelins and trypsin-like serine proteinases, in which the first cut during activation generates the mature N-terminus (Avilés et al., 1993; Khan and James, 1998). Subsequently, the pre-mature astacin variants, which are catalytically active (Yiallouros et al., 2002), carry out further autolytic cleavages, eventually giving rise to the competent N-terminus at A^1 (see chapter 'A buried N-terminus in mature astacins'). In matrix metalloproteinases, similar trimming yields the competent N-terminus needed to form a salt bridge with a conserved aspartate (Reinemer et al., 1994; Nagase, 1997). Finally, comparison of zymogenic and mature astacin further reveals that upon cleavage at G^{34P}-A^1, the main chain must undergo a 180° rotation around the Ψ main-chain angle of the new N-terminal residue to allow segment A^1-D^6 to insert into the molecular moiety and fulfill the plugging function described in the chapter 'A buried N-terminus in mature astacins' above.

Protein inhibitors

Astacins are not inhibited by tissue inhibitors of metalloproteinases (TIMPs). A natural inhibitor of astacin is the general protein scavenger α_2-macroglobulin (Stöcker et al., 1991a; Meier et al., 1994; Zhang et al., 2006; Marrero et al., 2012). However, α_2-macroglobulin will only entrap endopeptidases of limited size. Larger oligomeric astacins, such as the meprins, are not inhibited by this regulator of vascular and interstitial proteolysis. Interestingly, there is a fish astacin from carp head kidney (a hematopoietic organ) which circulates in the blood stream in complex with a specific protein inhibitor. Originally, this enzyme was termed 'nephrosin' (Hung et al., 1997) and its inhibitor 'nephrosin inhibitor' (Tsai et al., 2004). The nephrosin inhibitor is a homolog of fetuin, a large plasma protein with many functions. Fish fetuin, like its mammalian counterpart fetuin A, contains cystatin-like domains and is related to cystatin C-like inhibitors of cysteine cathepsins. More recently, it has been demonstrated that the plasma proteins fetuin A and cystatin C act as physiological inhibitors of human astacin proteases such as meprins and also block crayfish astacin (Hedrich et al., 2010).

Conclusions

Since the first report on astacins back in 1991 (Dumermuth et al., 1991), the family has expanded to several hundreds of animal and bacterial – though not plant or archeal – sequences (Sterchi et al., 2008). Six are found in humans and up to 40 in *C. elegans* (Möhrlen et al., 2003), and putative uncharacterized proteins from *Fusarium oxysporum* (UniProt entries

F9FJL4, F9FJL6, and F9FF60), *Phaeosphaeria nodorum* (Q0U0C2 and Q0UTK3), *Pyrenophora teres* (E3RUK5), *Pyrenophora tritici-repentis* (B2WI05), and *Unicarpus reesii* (C4JMI3) may represent the first fungal members. Through their degradatory potential, astacins are involved in embryonic development, tissue differentiation, and extracellular matrix assembly, and they are thus therapeutic targets (Mac Sweeney et al., 2008; Sterchi et al., 2008; Okada et al., 2010). Physiologically, they are regulated by zymogenicity and co-localizing protein inhibitors; however, to be able to act upon them from a therapeutic perspective, it will be necessary to design highly specific small-molecule inhibitors that target the catalytic moieties. This is where detailed knowledge of their three-dimensional structure, tentatively presented in this review, may prove helpful.

References

Ambort, D., Brellier, F., Becker-Pauly, C., Stöcker, W., Andrejevic-Blant, S., Chiquet, M., and Sterchi, E.E. (2010). Specific processing of tenascin-C by the metalloprotease meprin β neutralizes its inhibition of cell spreading. Matrix Biol. *29*, 31–42.

Apte, S.S. (2009). A disintegrin-like and metalloprotease (reprolysin-type) with thrombospondin type 1 motif (ADAMTS) superfamily: functions and mechanisms. J. Biol. Chem. *284*, 31493–31497.

Avilés, F.X., Vendrell, J., Guasch, A., Coll, M., and Huber, R. (1993). Advances in metallo-procarboxypeptidases. Emerging details on the inhibition mechanism and on the activation process. Eur. J. Biochem. *211*, 381–389.

Banerjee, S. and Bond, J. (2008). Prointerleukin-18 is activated by meprin β *in vitro* and *in vivo* in intestinal inflammation. J. Biol. Chem. *283*, 31371–31377.

Baumann, U. (2004). 152. Serralysin and related enzymes. In: A.J. Barrett, N.D. Rawlings, and J.F. Woessner Jr., eds. Handbook of Proteolytic Enzymes, Vol. 1 (London: Elsevier Academic Press), pp. 579–581.

Bayés, A., Fernández, D., Solà, M., Marrero, A., García-Piqué, S., Avilés, F.X., Vendrell, J., and Gomis-Rüth, F.X. (2007). Caught after the act: a human A-type metallocarboxypeptidase in a product complex with a cleaved hexapeptide. Biochemistry *46*, 6921–6930.

Becker-Pauly, C., Bruns, B.C., Damm, O., Schütte, A., Hammouti, K., Burmester, T., and Stöcker, W. (2009). News from an ancient world: two novel astacin metalloproteases from the horseshoe crab. J. Mol. Biol. *385*, 236–248.

Becker-Pauly, C., Barré, O., Schilling, O., Auf dem Keller, U., Ohler, A., Broder, C., Schutte, A., Kappelhoff, R., Stöcker, W., and Overall, C.M. (2011). Proteomic analyses reveal an acidic prime side specificity for the astacin metalloprotease family reflected by physiological substrates. Mol. Cell. Proteomics 10, M111.009233.

Beckmann, G. and Bork, P. (1993). An adhesive domain detected in functionally diverse receptors. Trends Biochem. Sci. 18, 40–41.

Bode, W. and Huber, R. (1978). Crystal structure analysis and refinement of two variants of trigonal trypsinogen: trigonal trypsin and PEG (polyethylene glycol) trypsinogen and their comparison with orthorhombic trypsin and trigonal trypsinogen. FEBS Lett. 90, 265–269.

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.

Bode, W., Gomis-Rüth, F.X., and Stöcker, W. (1993). Astacins, serralysins, 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.

Boldt, H.B., Overgaard, M.T., Laursen, L.S., Weyer, K., Sottrup-Jensen, L., and Oxvig, C. (2001). Mutational analysis of the proteolytic domain of pregnancy-associated plasma protein-A (PAPP-A): classification as a metzincin. Biochem. J. 358, 359–367.

Bond, J.S. and Beynon, R.J. (1995). The astacin family of metalloendopeptidases. Protein Sci. 4, 1247–1261.

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, 498–506.

Bork, P. and Beckmann, G. (1993). The CUB domain. A widespread module in developmentally regulated proteins. J. Mol. Biol. 231, 539–545.

Butler, G.S., Tam, E.M., and Overall, C.M. (2004). The canonical methionine 392 of matrix metalloproteinase 2 (gelatinase A) is not required for catalytic efficiency or structural integrity: probing the role of the methionine-turn in the metzincin metalloprotease superfamily. J. Biol. Chem. 279, 15615–15620.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Bryan Arendall, W. 3rd, Snoeyink, J., Richardson, J.S., et al. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 35 (Web Server issue), W375–W383.

Dumermuth, E., Sterchi, E.E., Jiang, W.P., Wolz, R.L., Bond, J.S., Flannery, A.V., and Beynon, R.J. (1991). The astacin family of metalloendopeptidases. J. Biol. Chem. 266, 21381–21385.

Elaroussi, M.A. and DeLuca, H.F. (1994). A new member to the astacin family of metalloendopeptidases: a novel 1,25-dihydroxyvitamin D-3-stimulated mRNA from chorioallantoic membrane of quail. Biochim. Biophys. Acta 1217, 1–8.

Fan, T.J. and Katagiri, C. (2001). Properties of the hatching enzyme from Xenopus laevis. Eur. J. Biochem. 268, 4892–4898.

Fehlhammer, H., Bode, W., and Huber, R. (1977). Crystal structure of bovine trypsinogen at 1.8 Å resolution. II. Crystallographic refinement, refined crystal structure and comparison with bovine trypsin. J. Mol. Biol. 111, 415–438.

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, 49835–49841.

Ge, G. and Greenspan, D.S. (2006). Developmental roles of the BMP1/TLD metalloproteinases. Birth Defects Res., Part C 78, 47–68.

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, 796–803.

Gomis-Rüth, F.X. (2003). Structural aspects of the metzincin clan of metalloendopeptidases. Mol. Biotechnol. 24, 157–202.

Gomis-Rüth, F.X. (2008). Structure and mechanism of metallocarboxypeptidases. Crit. Rev. Biochem. Mol. Biol. 43, 319–345.

Gomis-Rüth, F.X. (2009). Catalytic domain architecture of metzincin metalloproteases. J. Biol. Chem. 284, 15353–15357.

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, 945–968.

Gomis-Rüth, F.X., Nar, H., Grams, F., Yallouros, I., Küsthardt, U., Zwilling, R., Bode, W., and Stöcker, W. (1994). Crystal structures, spectroscopic features and catalytic properties of cobalt(II)-, copper(II)-, nickel(II)-, and mercury(II)-derivatives of the zinc-endopeptidase astacin. A correlation of structure and proteolytic activity. J. Biol. Chem. 269, 17111–17117.

Gomis-Rüth, F.X., Botelho, T.O., and Bode, W. (2012). A standard orientation for metallopeptidases. Biochim. Biophys. Acta 1824, 157–163.

Goulas, T., Arolas, J.L., and Gomis-Rüth, F.X. (2010). Structure, function and latency regulation of a bacterial enterotoxin potentially derived from a mammalian adamalysin/ADAM xenolog. Proc. Natl. Acad. Sci. USA 108, 1856–1861.

Grams, F., Dive, V., Yiotakis, A., Yiallouros, I., Vassiliou, S., Zwilling, R., Bode, W., and Stöcker, W. (1996). Structure of astacin with a transition-state analogue inhibitor. Nat. Struct. Biol. 3, 671–675.

Guevara, T., Yiallouros, I., Kappelhoff, R., Bissdorf, S., Stöcker, W., and Gomis-Rüth, F.X. (2010). Proenzyme structure and activation of astacin metallopeptidase. J. Biol. Chem. 285, 13958–13965.

Hahn, D., Pischitzis, A., Roesmann, S., Hansen, M.K., Leuenberger, B., Luginbuehl, U., and Sterchi, E.E. (2003). Phorbol 12-myristate 13-acetate-induced ectodomain shedding and phosphorylation of the human meprin β metalloprotease. J. Biol. Chem. 278, 42829–42839.

Hedrich, J., Lottaz, D., Meyer, K., Yiallouros, I., Jahnen-Dechent, W., Stöcker, W., and Becker-Pauly, C. (2010). Fetuin-A and cystatin C are endogenous inhibitors of human meprin metalloproteases. Biochemistry 49, 8599–8607.

Hege, T. and Baumann, U. (2001). The conserved methionine residue of the metzincins: a site-directed mutagenesis study. J. Mol. Biol. 314, 181–186.

Herzog, C., Kaushal, G.P., and Haun, R.S. (2005). Generation of biologically active interleukin-1β by meprin B. Cytokine 31, 394–403.

Hintze, V., Höwel, M., Wermter, C., Grosse Berkhoff, E., Becker-Pauly, C., Beermann, B., Yiallouros, I., and Stöcker, W. (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, 6741–6748.

Holland, D.R., Tronrud, D.E., Pley, H.W., Flaherty, K.M., Stark, W., Jansonius, J.N., McKay, D.B., and Matthews, B.W. (1992). Structural comparison suggests that thermolysin and related neutral proteases undergo hinge-bending motion during catalysis. Biochemistry 31, 11310–11316.

Holley, S., Neul, J., Attisano, L., Wrana, J., Sasai, Y., O'Connor, M., De Robertis, E., and Ferguson, E. (1996). The Xenopus

dorsalizing factor noggin ventralizes *Drosophila* embryos by preventing DPP from activating its receptor. Cell *86*, 607–617.

Hopkins, D.R., Keles, S., and Greenspan, D.S. (2007). The bone morphogenetic protein 1/Tolloid-like metalloproteinases. Matrix Biol. *26*, 508–523.

Huber, R. and Bode, W. (1986). Structural basis for the activation and action of trypsin. Acc. Chem. Res. *11*, 114–122.

Hung, C.H., Huang, H.R., Huang, C.J., Huang, F.L., and Chang, G.D. (1997). Purification and cloning of carp nephrosin, a secreted zinc endopeptidase of the astacin family. J. Biol. Chem. *272*, 13772–13778.

Jefferson, T., Čaušević, M., Auf dem Keller, U., Schilling, O., Isbert, S., Geyer, R., Maier, W., Tschickardt, S., Jumpertz, T., Weggen, S., et al. (2011). The metalloprotease meprin β generates non-toxic N-terminal amyloid precursor protein fragments *in vivo*. J. Biol. Chem. *286*, 27741–27750.

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.

Khan, A.R. and James, M.N. (1998). Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. Protein Sci. *7*, 815–836.

Kronenberg, D., Bruns, B.C., Moali, C., Vadon-Le Goff, S., Sterchi, E.E., Traupe, H., Böhm, M., Hulmes, D.J., Stöcker, W., and Becker-Pauly, C. (2010). Processing of procollagen III by meprins: new players in extracellular matrix assembly? J. Invest. Dermatol. *130*, 2727–2735.

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*, 147–163.

Li, S.W., Sieron, A.L., Fertala, A., Hojima, Y., Arnold, W.V., and Prockop, D.J. (1996). The C-proteinase that processes procollagens to fibrillar collagens is identical to the protein previously identified as bone morphogenic protein-1. Proc. Natl. Acad. Sci. USA *93*, 5127–5130.

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*, 27–37.

Mac Sweeney, A., Gil-Parrado, S., Vinzenz, D., Bernardi, A., Hein, A., Bodendorf, U., Erbel, P., Logel, C., and Gerhartz, B. (2008). Structural basis for the substrate specificity of bone morphogenetic protein 1/tolloid-like metalloproteases. J. Mol. Biol. *384*, 228–239.

Marrero, A., Duquerroy, S., Trapani, S., Goulas, T., Guevara, T., Andersen, G.R., Navaza, J., Sottrup-Jensen, L., and Gomis-Rüth, F.X. (2012). The crystal structure of human α₂-macroglobulin reveals a unique molecular cage. Angew. Chem. Int. Ed. *51*, 3340–3344.

Matthews, B.W. (1988). Structural basis of the action of thermolysin and related zinc peptidases. Acc. Chem. Res. *21*, 333–340.

Matthews, B.W., Jansonius, J.N., Colman, P.M., Schoenborn, B.P., and Dupourque, D. (1972). Three-dimensional structure of thermolysin. Nature *238*, 37–41.

McCall, K.A., Huang, C., and Fierke, C.A. (2000). Function and mechanism of zinc metalloenzymes. J. Nutr. *130*, 1437S–1446S.

Meier, U.C., Boetzel, J., Kellermann, J., Mann, K., Billich, A., Stöcker, W., and Schramm, H.J. (1994). The cleavage of the bait region of α₂-macroglobulin by human immunodeficiency virus proteinases and by astacin. Ann. NY Acad. Sci. *737*, 431–433.

Möhrlen, F., Hutter, H., and Zwilling, R. (2003). The astacin protein family in *Caenorhabditis elegans*. Eur. J. Biochem. *270*, 4909–4920.

Nagase, H. (1997). Activation mechanisms of matrix metalloproteinases. Biol. Chem. *378*, 151–160.

Oberholzer, A.E., Bumann, M., Hege, T., Russo, S., and Baumann, U. (2009). Metzincin's canonical methionine is responsible for the structural integrity of the zinc-binding site. Biol. Chem. *390*, 875–881.

Ohler, A., Debela, M., Wagner, S., Magdolen, V., and Becker-Pauly, C. (2010). Analyzing the protease web in skin: meprin metalloproteases are activated specifically by KLK4, 5 and 8 *vice versa* leading to processing of proKLK7 thereby triggering its activation. Biol. Chem. *391*, 455–460.

Okada, A., Sano, K., Nagata, K., Yasumasu, S., Ohtsuka, J., Yamamura, A., Kubota, K., Iuchi, I., and Tanokura, M. (2010). Crystal structure of zebrafish hatching enzyme 1 from the zebrafish *Danio rerio*. J. Mol. Biol. *402*, 865–878.

Pauptit, R.A., Karlsson, R., Picot, D., Jenkins, J.A., Niklaus-Reimer, A.S., and Jansonius, J.N. (1988). Crystal structure of neutral protease from *Bacillus cereus* refined at 3.0 Å resolution and comparison with the homologous but more thermostable enzyme thermolysin. J. Mol. Biol. *199*, 525–537.

Pérez, L., Kerrigan, J.E., Li, X., and Fan, H. (2007). Substitution of methionine 435 with leucine, isoleucine, and serine in tumor necrosis factor α converting enzyme inactivates ectodomain shedding activity. Biochem. Cell Biol. *85*, 141–149.

Pieper, M., Betz, M., Budisa, N., Gomis-Rüth, F.X., Bode, W., and Tschesche, H. (1997). Expression, purification, characterization, and X-ray analysis of selenomethionine 215 variant of leukocyte collagenase. J. Protein Chem. *16*, 637–650.

Quesada, V., Sánchez, L.M., Alvarez, J., and López-Otín, C. (2004). Identification and characterization of human and mouse ovastacin: a novel metalloproteinase similar to hatching enzymes from arthropods, birds, amphibians, and fish. J. Biol. Chem. *279*, 26627–26634.

Rawlings, N.D., Barrett, A.J., and Bateman, A. (2010). MEROPS: the peptidase database. Nucleic Acids Res. *38*, D227–D233.

Reinemer, P., Grams, F., Huber, R., Kleine, T., Schnierer, S., Piper, M., Tschesche, H., and Bode, W. (1994). Structural implications for the role of the N-terminus in the 'superactivation' of collagenases. A crystallographic study. FEBS Lett. *338*, 227–233.

Reyda, S., Jacob, E., Zwilling, R., and Stöcker, W. (1999). cDNA cloning, bacterial expression, *in vitro* renaturation and affinity purification of the zinc endopeptidase astacin. Biochem. J. *344*, 851–857.

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*, 769–786.

Rothe, M., Wong, S.C., Henzel, W.J., and Goeddel, D.V. (1994). A novel family of putative signal transducers associated with the cytoplasmic domain of the 75 kDa tumor necrosis factor receptor. Cell *78*, 681–692.

Sachdev, M., Mandal, A., Mulders, S., Digilio, L.C., Panneerdoss, S., Suryavathi, V., Pires, E., Klotz, K.L., Hermens, L., Herrero, M.B., et al. (2012). Oocyte specific oolemmal SAS1B involved in sperm binding through intra-acrosomal SLLP1 during fertilization. Dev. Biol. *363*, 40–51.

Sato, S.M. and Sargent, T.D. (1990). Molecular approach to dorsoanterior development in *Xenopus laevis*. Dev. Biol. *137*, 135–141.

Schütte, A., Hedrich, J., Stöcker, W., and Becker-Pauly, C. (2010). Let it flow: Morpholino knockdown in zebrafish embryos reveals

a pro-angiogenic effect of the metalloprotease meprin α2. PLoS ONE 5, e8835.

Shibata, Y., Iwamatsu, T., Oba, Y., Kobayashi, D., Tanaka, M., Nagahama, Y., Suzuki, N., and Yoshikuni, M. (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, 8349–8354.

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.

Sieron, A.L., Tretiakova, A., Jameson, B.A., Segall, M.L., Lund-Katz, S., Khan, M.T., Li, S.W., and Stöcker, W. (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, 3231–3239.

Sonneborn, H.H., Zwilling, R., and Pfleiderer, G. (1969). [Evolution of endopeptidases. X. Cleavage specificity of low molecular weight protease from *Astacus leptodactylus* Esch]. Hoppe-Seylers Z. Physiol. Chem. 350, 1097–1102.

Sterchi, E.E., Stöcker, W., and Bond, J.S. (2008). Meprins, membrane-bound and secreted astacin metalloproteinases. Mol. Aspects Med. 29, 309–328.

Stöcker, W. and Bode, W. (1995). Structural features of a superfamily of zinc-endopeptidases: the metzincins. Curr. Opin. Struct. Biol. 5, 383–390.

Stöcker, W., Wolz, R.L., Zwilling, R., Strydom, D.J., and Auld, D. (1988). *Astacus* protease, a zinc metalloenzyme. Biochemistry 27, 5026–5032.

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.

Stöcker, W., Breit, S., Sottrup-Jensen, L., and Zwilling, R. (1991a). α2-Macroglobulin from hemolymph of the freshwater crayfish *Astacus astacus*. Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol. 98, 501–509.

Stöcker, W., Sauer, B., and Zwilling, R. (1991b). Kinetics of nitroanilide cleavage by astacin. Biol. Chem. Hoppe-Seyler 372, 385–392.

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.

Stöcker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Rüth, F.X., McKay, D.B., and Bode, W. (1995). The metzincins – topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases. Protein Sci. 4, 823–840.

Takeda, S. (2009). Three-dimensional domain architecture of the ADAM family proteinases. Semin. Cell Dev. Biol. 20, 146–152.

Takeda, S., Takeya, H., and Iwanaga, S. (2012). Snake venom metalloproteinases: structure, function and relevance to the mammalian ADAM/ADAMTS family proteins. Biochim. Biophys. Acta 1824, 164–176.

Tallant, C., García-Castellanos, R., Seco, J., Baumann, U., and Gomis-Rüth, F.X. (2006). Molecular analysis of ulilysin, the structural prototype of a new family of metzincin metalloproteases. J. Biol. Chem. 281, 17920–17928.

Tallant, C., García-Castellanos, R., Baumann, U., and Gomis-Rüth, F.X. (2010a). On the relevance of the Met-turn methionine in metzincins. J. Biol. Chem. 285, 13951–13957.

Tallant, C., Marrero, A., and Gomis-Rüth, F.X. (2010b). Matrix metalloproteinases: fold and function of their catalytic domains. Biochim. Biophys. Acta, Mol. Cell Res. 1803, 20–28.

Thayer, M.M., Flaherty, K.M., and McKay, D.B. (1991). Three-dimensional structure of the elastase of *Pseudomonas aeruginosa* at 1.5-Å resolution. J. Biol. Chem. 266, 2864–2871.

Titani, K., Torff, H.J., Hormel, S., Kumar, S., Walsh, K.A., Rödl, J., Neurath, H., and Zwilling, R. (1987). Amino acid sequence of a unique protease from the crayfish *Astacus fluviatilis*. Biochemistry 26, 222–226.

Tsai, P.L., Chen, C.H., Huang, C.J., Chou, C.M., and Chang, G.D. (2004). Purification and cloning of an endogenous protein inhibitor of carp nephrosin, an astacin metalloproteinase. J. Biol. Chem. 279, 11146–11155.

Vazeille, E., Bringer, M.A., Gardarin, A., Chambon, C., Becker-Pauly, C., Pender, S.L., Jakob, C., Müller, S., Lottaz, D., and Darfeuille-Michaud, A. (2011). Role of meprins to protect ileal mucosa of Crohn's disease patients from colonization by adherent-invasive *E. coli*. PLoS ONE 6, e21199.

Walasek, P. and Honek, J.F. (2005). Nonnatural amino acid incorporation into the methionine 214 position of the metzincin *Pseudomonas aeruginosa* alkaline protease. BMC Biochem. 6, 21.

Waltersperger, S., Widmer, C., Wang, M., and Baumann, U. (2010). Crystal structure of archaemetzincin AmzA from *Methanopyrus kandleri* at 1.5 Å resolution. Proteins 78, 2720–2723.

Wermter, C., Höwel, M., Hintze, V., Bombosch, B., Aufenvenne, K., Yiallouros, I., and Stöcker, W. (2007). The protease domain of procollagen C-proteinase (BMP1) lacks substrate selectivity, which is conferred by non-proteolytic domains. Biol. Chem. 388, 513–521.

Wolfsberg, T., Bazan, J., Blobel, C., Myles, D., Primakoff, P., and White, J. (1993). The precursor region of a protein active in sperm-egg fusion contains a metalloprotease and a disintegrin domain: structural, functional, and evolutionary implications. Proc. Natl. Acad. Sci. USA 90, 10783–10787.

Wozney, J.M., Rosen, V., Celeste, A.J., Mitsock, L.M., Whitters, M.J., Kriz, R.W., Hewick, R.M., and Wang, E.A. (1988). Novel regulators of bone formation: molecular clones and activities. Science 242, 1528–1534.

Yan, L., Fei, K., Zhang, J., Dexter, S., and Sarras, M.P. Jr. (2000a). Identification and characterization of hydra metalloproteinase 2 (HMP2): a meprin-like astacin metalloproteinase that functions in foot morphogenesis. Development 127, 129–141.

Yan, L., Leontovich, A., Fei, K., and Sarras, M.P. Jr. (2000b). Hydra metalloproteinase 1: a secreted astacin metalloproteinase whose apical axis expression is differentially regulated during head regeneration. Dev. Biol. 219, 115–128.

Yasumasu, S., Yamada, K., Akasaka, K., Mitsunaga, K., Iuchi, I., Shimada, H., and Yamagami, K. (1992). Isolation of cDNAs for LCE and HCE, two constituent proteases of the hatching enzyme of *Oryzias latipes*, and concurrent expression of their mRNAs during development. Dev. Biol. 153, 250–258.

Yasumasu, S., Shimada, H., Inohaya, K., Yamazaki, K., Iuchi, I., Yasumasu, I., and Yamagami, K. (1996). Different exon-intron organizations of the genes for two astacin-like proteases, high choriolytic enzyme (choriolysin H) and low choriolytic enzyme (choriolysin L), the constituents of the fish hatching enzyme. Eur. J. Biochem. 237, 752–758.

Yiallouros, I., Grosse-Berkhoff, E., and Stöcker, W. (2000). The roles of Glu93 and Tyr149 in astacin-like zinc peptidases. FEBS Lett. 484, 224–228.

Yiallouros, I., Kappelhoff, R., Schilling, O., Wegmann, F., Helms, M.W., Auge, A., Brachtendorf, G., Berkhoff, E.G., Beermann,

B., Hinz, H.J., et al. (2002). Activation mechanism of pro-asta-cin: role of the pro-peptide, tryptic and autoproteolytic cleavage and importance of precise amino-terminal processing. J. Mol. Biol. *324*, 237–246.

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*, 24242–24252.

Zhang, Y., Ge, G., and Greenspan, D.S. (2006). Inhibition of bone morphogenetic protein 1 by native and altered forms of α_2-macroglobulin. J. Biol. Chem. *281*, 39096–39104.

F. Xavier Gomis-Rüth performed his PhD on protein crystallography under supervision of Wolfram Bode in the department of Robert Huber at the Max Planck Institute of Biochemistry (MPIB) in Martinsried (Germany) in 1992. Post-doctoral stays followed at the Autonomous University of Barcelona, the MPIB, and the Molecular Biology Institute of Barcelona (IBMB) of the Spanish Research Council (CSIC), where he became a tenured research scientist in 2000. Since 2008, he has been a Research Professor at CSIC. He is the PI of the Proteolysis Lab at the Department of Structural Biology of the IBMB and author of over 100 publications in the field.

Sergio Trillo-Muyo studied human biology at the Pompeu Fabra University of Barcelona, Spain, and obtained his BSc in 2006. In 2008, he obtained an MSc in Biochemistry and Molecular Biology at the Autonomous University of Barcelona in Bellaterra, Spain. He is currently a PhD student in the Proteolysis Lab at the Department of Structural Biology of the Molecular Biology Institute of Barcelona, which is part of the Spanish Research Council (CSIC).

Walter Stöcker studied chemistry and biology at the University of Munich, Germany. After receiving his doctoral degree in Munich in 1984, he was a scientific assistant at the Institute of Zoology, University of Heidelberg, Germany, and did research as a visiting scientist at the Harvard Medical School, Boston, MA, USA, and the Max Planck Institute of Biochemistry, Martinsried, Germany. He received his habilitation in biology from the University of Heidelberg in 1992, and became Associate Professor of Zoology at the University of Münster, Germany in 1996. Since 2003, he has been a full professor and chair at the University of Mainz, Germany. His research focuses on the structure and function of astacin-like zinc peptidases and their interactions in the proteolytic web.