

Evolution of astacin-like metalloproteases in animals and their function in development

Frank Möhrle^{a,*}, Melanie Maniura^b, Günter Plickert^b, Marcus Frohme^c, and Uri Frank^{a,*}

^aDepartment of Zoology and Martin Ryan Marine Science Institute, National University of Ireland, Galway, Ireland.

^bInstitute of Zoology, University of Cologne, Germany

^cDepartment of Functional Genome Analysis, German Cancer Research Center, Heidelberg, Germany

*Authors for correspondence (email: moehrlen@uni-hd.de; uri.frank@nuigalway.ie)

SUMMARY Astacin-like metalloproteases are ubiquitous in the animal kingdom but their phylogenetic relationships and ancient functions within the Metazoa are unclear. We have cloned and characterized four astacin-like cDNAs from the marine hydroid *Hydractinia echinata* and performed a database search for related genes in the draft genome sequence of the sea anemone *Nematostella vectensis*. These sequences and those of higher animals' astacins were subjected to phylogenetic analysis revealing five clusters within the Eumetazoa. The bone morphogenetic protein-1/tolloid-like astacins were represented in all eumetazoan phyla studied. The meprins were only found in vertebrates and

cnidarians. Two clusters were taxon-specific, and one cluster represented astacins, which probably evolved after the split of the Cnidaria. Interestingly, grouping of astacins according to the protease catalytic domain alone resulted in clusters of proteins with similar overall domain architecture. The *Hydractinia* astacins were expressed in distinct cells during metamorphosis and some also during wound healing. Previously characterized cnidarian astacins also act during development. Based on our phylogeny, however, we propose that the developmental function of most of them is not homologous to the developmental function assigned to higher animals' astacins.

INTRODUCTION

Members of the astacin family of zinc endopeptidases, a subclass of the metzincins, are found throughout the animal kingdom and also in some bacterial species (Bond and Beynon 1995; Stöcker et al. 1995). Astacins exert their function in the extracellular space, occurring either as membrane bound—or secreted enzymes (for an overview see Zwilling and Stöcker 1997; Barrett et al. 2004). They can be classified according to their structures into three main groups, which roughly also correspond to their physiological function. Meprins are membrane bound enzymes, found in the small intestine and kidney of mammals. They are believed to be involved in the processing of biologically active peptides (Dumermuth et al. 1991; Bankus and Bond 1996; Köhler et al. 2000). The bone morphogenetic protein (BMP)-1/tolloids comprise morphogenetically active proteins, found in many vertebrates and invertebrates. BMP-1, also termed procollagen C proteinase, is an enzyme critical for the assembly of collagen fibers (Reddi 1996). Tolloid-like proteins of vertebrates and tolloid of *Drosophila* are involved in the regulation of the transforming growth factor BMP-4 and the SOG/DPP complex, respectively, and thus in the specification of ventral body struc-

tures (Suzuki et al. 1996; Blader et al. 1997; Mullins 1998). Both meprins and BMP-1/tolloids are characterized by long, complex C-terminal extensions adjacent to the catalytic protease domain, which presumably define the time and place of their activities. Meprins and BMP-1/tolloids are also characterized by very specific substrate preferences of their catalytic domains. Finally, astacin-like enzymes have also been shown to be digestive enzymes and to be engaged in the breakdown of the egg envelope during hatching of nematodes, crustaceans, fishes, frogs, and birds (Yasumasu et al. 1992; Hishida et al. 1996; Geier and Zwilling 1998; Shibata et al. 2000; Fan and Katagiri 2001). This group of astacins, however, is rather diverse, having a broad substrate preference. Unlike other astacins, they are typically comprised of only the catalytic domain without other significant C-terminal segments.

Interestingly, classification of vertebrates, *Drosophila* and *Caenorhabditis elegans* astacins according to the sequence of the catalytic domain alone matches the classification according to the C-terminal domain architecture (Möhrle et al. 2003). This point suggests that the diversification of the main astacin groups predates the diversification of these taxa. It may also imply that the functions of these main astacin groups have been conserved too. It is, however, unclear at

which point in metazoan evolution this diversification occurred and what the functions of early astacins were.

A partial answer to these questions has been provided by previous work on lower invertebrates' astacins. Also in cnidarians, astacins have been suggested to play an important role in developmental processes. *Hydra* metalloproteinase 1 (HMP-1) is believed to be involved in head regeneration and transdifferentiation of tentacle battery cells (Yan et al. 2000b). *Hydra* metalloproteinase 2 (HMP-2) was found to be active during foot morphogenesis (Yan et al. 2000a). The foot activator responsive metalloprotease (FARM-1), the third known *Hydra* astacin, was identified as a possible target gene for foot activator peptides (Kumpfmüller et al. 1999). Finally, metalloproteinase 1 (PMP-1) from the hydroid *Podocoryne carnea* shows a putative dual role in medusae bud development and in food digestion (Pan et al. 1998). In contrast to vertebrate or *Drosophila* astacins, however, cnidarian morphogenetic active astacins are mostly characterized by simple C-termini. The phylogenetic relationship of these astacins is unclear.

This investigation therefore was prompted by the question of how many different functional and evolutionary clusters astacins represent in the ancient eumetazoan phylum, the Cnidaria. For our study we chose the colonial marine hydroid *Hydractinia echinata* as a model organism because it offers an easy, year-round access to embryonic development and metamorphosis of the planula larvae; studies which are not feasible with *Hydra*. We have cloned four astacin-like genes from *Hydractinia*, studied their role in development and metamorphosis, and analyzed their phylogenetic relationships in the animal realm.

MATERIALS AND METHODS

Animal cultures

H. echinata colonies were cultured in artificial seawater as described (Frank et al. 2001). Fertilized eggs were collected daily about 2 h after light-induced spawning and allowed to develop in sterile seawater for 3 days until the planula stage was reached. They were then induced to undergo metamorphosis by a 3 h incubation in 100 mM CsCl in seawater (Müller and Leitz 2002). Following metamorphosis induction, animals were positioned on glass coverslips to complete metamorphosis, which is schematically shown in Fig. 1. Metamorphosed animals were fed every other day with 3-day-old brine shrimps nauplii.

Cloning of astacin genes

Two cDNA fragments showing high-sequence similarity to known astacins were identified during an in situ hybridization screen, aimed at identifying genes up-regulated during metamorphosis. Two additional cDNAs were found in our *Hydractinia* expressed sequence tags (EST) database. The genes were termed *HEA1-4* for *H. echinata* astacins. Full-length cDNAs were obtained by rapid

amplification of cDNA ends-polymerase chain reaction (RACE, PCR) using the SMART RACE kit (Clontech, Mountain View, CA, USA) according to the manufacturer's protocol. RACE primers used were F 5'-CAA CGC AAT CAA CGA AG-3'+R 5'AAA TCA GCT TGG CCA AGT G3', F 5'-AAC TGC AGA GGG CAT ATG GAA AAC TGG AAA TGC TT-3'+R 5'-TAG CGG CCG CGT GAA GAG CAG CCG CCT CCT TTA TAG-3', F5'—CGG GGC TGC TAG TAC TTT AGG AGA ACC-3'+R 5'-CAC CAC TTT CCT AAC ACG CCA CAA AAT-3', and F 5'-GCC GAA GCT TGC CAC TTT TCT TTG TTA GTC CTC-3'+R 5'-GCC TCT AGA CAC AAA ATG AAG CTC GTG TTT A-3' for *HEA1-HEA4*, respectively. PCR products were cloned into pBlueScript (Stratagene, Amsterdam, the Netherlands) and sequenced from both ends with vector-specific primers.

Sequence analysis

Nucleotide and protein sequences were analyzed using BLAST at the NCBI website. To identify astacin-like genes in the genome of *Nematostella vectensis*, we used representative cnidaria and vertebrate proteins, or their conserved domains according to the PFAM and PRINTS database, as queries for BLAST searches. The occurrence of signal peptides was predicted using SignalP V2.0.b2 at <http://www.cbs.dtu.dk/services/SignalP-2.0/>. For identification and annotation of protein domains and the analysis of domain architectures the tools of the SMART, PFAM, ProDom and INTERPRO protein domain databases were used.

For phylogenetic studies protein sequences of the catalytic domain of astacins, representing the main metazoan clades, and all astacin subgroups, were aligned using CLUSTAL W and imported into GeneDoc for further manipulation (data available upon request). An alignment of several astacins is depicted in Fig. 2. Phylogenetic inference was carried out using the Bayesian phylogenetic analysis and the MrBayes 3.0beta4 program (Huelsenbeck and Ronquist 2001). The WAG model with γ distribution of substitution rates was applied. Prior probabilities for all trees were equal, the starting trees were random and tree sampling was performed every 10th generation. Metropolis-coupled Markov chain Monte Carlo sampling was performed with one cold and three heated chains that were run for 50,000 generations. Posterior probabilities were estimated on 3,000 trees (burnin = 2,000). The tree presented here was visualized using TreeView (<http://taxonomy.zoology.gla.ac.uk/rod/rod.html>)

In situ hybridization

The same primers used for RACE were also used to PCR amplify fragments of 277, 223, 270, and 217 bp for *HEA1-4*, respectively. PCR products were cloned into pGEM-T (Promega, Mannheim, Germany) or pBluescript (Stratagene) vectors. The vectors were linearized by restriction digestion to generate templates for in vitro RNA synthesis. Sense and anti-sense RNA probes were synthesized using T7 and SP6 RNA polymerases (for pGEM-T), or T3 and T7 (for pBluescript) with digoxigenin labeled nucleotides (Roche, Mannheim, Germany). In situ hybridization was carried out as described (Gajewsky et al. 1996) on samples representing the entire life cycle of the animal. At an early stage of the project we realized that a strong signal appeared in recovering injuries. We have then designed a new series of experiments in which various

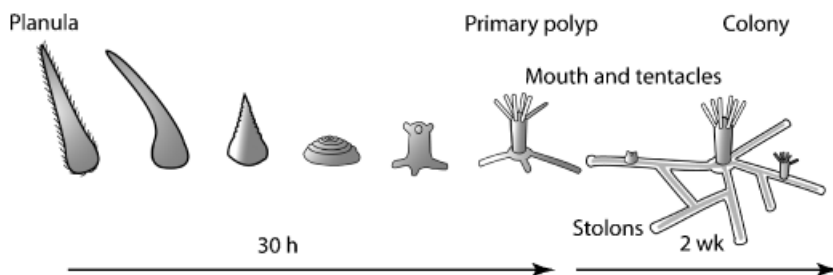


Fig. 1. Schematic representation of metamorphosis and its time frame in *Hydractinia echinata*.

tissue lesions were inflicted to polyps. Animals were fixed in 2–4 h intervals following injuries starting immediately after injury (<5 min) and up to 36 h.

RESULTS

Hydractinia astacins cDNAs

Four full-length cDNA clones, showing significant sequence similarity to astacin family members from other organisms, were obtained following 5' and 3' RACEs. The nucleotide sequences have been deposited in EMBL-Bank under accession numbers AM181342 (*HEA1*), AM181343 (*HEA2*), AM181344 (*HEA3*), and AM181345 (*HEA4*). The putative *HEA1* was a 137 amino acids protein. It contained a signal peptide at its N-terminal with a most likely cleavage site after position 18. The astacin domain was at position 19–137. This domain, however, is not complete. Alignment of *HEA1*'s catalytic domain with those of other astacins revealed that several critical amino-acid residues at the active center of the enzyme are missing (Fig. 2).

The cDNA of *HEA2* contained an open reading frame (ORF) encoding a protein of 289 amino-acid residues. It had a signal peptide at the N-terminus with a predicted cleavage site after position 19. A complete astacin domain was identified at position 50–246 and a ShKT domain (ShK toxin domain, SMART accession number: SM0254) at position 250–289.

The predicted gene product of *HEA3* comprised 304 amino-acid residues. It contained an astacin domain (positions 19–219) and two N-terminal ShTK domains in positions 230–263, and 265–304, respectively. This protein formed an exception to the other three *Hydractinia* astacins in having no detectable signal peptide.

HEA4 encoded a protein of 339 amino-acid residues. It was the longest among the four identified astacins. A signal peptide was predicted at the N-terminus with the most likely cleavage site between position 16 and 17. *HEA4* also contained one astacin domain (position 65–261), a disordered/unstructured region, as predicted by the DisEMBL HOT-LOOPS method and a single N-terminal ShTK domain at position 304–337.

Database search

Searching the draft genome sequence of the sea anemone *N. vectensis* we have identified a BMP-1/tolloid ortholog. The predicted protein contained all typical BMP-1/tolloid domains and the catalytic domain showed very high similarity to human BMP-1 (Fig. 2). The *Nematostella* genome sequence was also found to encode a meprin-like protein and several other astacins. However, the nucleotide sequence quality of the genomic contigs was not sufficient to reliably predict the amino-acid sequence of the entire processed enzymes, and these proteins could therefore not be included in the analysis.

Phylogenetic analysis

Using the Bayesian phylogenetic method, we have constructed a phylogenetic tree for the astacin protein family in the Eumetazoa (Fig. 3). Five main clusters were identified, two of which were taxon-specific: (1) the BMP-1/tolloid group, (2) the meprins, (3) *C. elegans*-specific astacins, (4) Cnidaria-specific astacins and (5) the hatching enzymes-like astacins. Astacins from cnidarians were represented in clusters #1 (BMP-1/tolloids), #2 (meprins) and #4 (Cnidaria-specific). Similar to the results of a previous study (Möhrlen et al. 2003), also here the catalytic domain-based analysis revealed clusters of astacins with similar C-terminal domain architecture. *Hydra* HMP2, which possesses a MAM domain, similar to all meprins, also clustered in the meprin group according to similarity in the astacin catalytic domain. The predicted *Nematostella* BMP-1/tolloid homolog clustered within the BMP-1/tolloid group.

In situ hybridization

The four *Hydractinia* astacins were mainly expressed during metamorphosis (Fig. 4). Their mRNA levels, as indicated by intensity of the chromogenic substrate reaction, were generally low in other developmental stages of intact animals, barely detected by in situ hybridization with only *HEA4* as a clear exception. Expression of *HEA1* started in just a few cells during early metamorphosis. Highest mRNA levels were detected at >16 h post-metamorphosis induction. Expression decreased toward the completion of metamorphosis and establishment of the primary polyp. The gene was expressed

A BMP1/Tolloids

```

*           20           *           40           *           60           *
HS_BMP1 : AATSRPERVWPDGVIPFVIGGNETGSRQRAVFRQAMRHWEKHTCVTFLEFTEDEDSYIVFTYRPGCCSYVG : 70
Nv_BMP-1 : AATAFVGRLLWPGGVIPFEIADBEFIDGQKATILQAMRHWFENFTCLSFIETHEEEDFIYFHKGRGCRGVS : 70

80           *           100          *           120          *           140
HS_BMP1 : RRGGGPOAISIGKNCDFEIVVHELGHVVGFWHEHTRPDRDRHVSIVRENIQPGQEYNFLKMEPQEVESL : 140
Nv_BMP-1 : RWPNTRSLLLTLGKWCDFKVVVHELGHITIGFWHEQNRPRDRDKYVKVMLEENVIPGEGHNFCKLTKAEIDSR : 140

*           160          *           180          *           200
HS_BMP1 : GETYDFDSIMHYARNTFSRGIIFLDITVPKYEVNGVKKPPIGQRTRLSKGDIACARKLYKCP : 200
Nv_BMP-1 : GEEYDYKSIMHYGKNTFATSLNHVTILPKRSGD-MVPAIGQREHLSLGDIPROTIKMYRCP : 199
    
```

B Meprins

```

*           20           *           40           *           60           *
HS_MEPA : NGLRDPNTRWTFP--IPYIILADNLGLNAKG--AIIYAFEMFRLKSCVDFKPYEGESSYIIFQCFDGCWS : 65
HV_HMP2 : AALSNSILLWLPNNKVVPSITKQLENTAEATFGLMAAFREWEERSCLTFKRRTDEKDYIBEFQCGSGCWS : 70

80           *           100          *           120          *           140
HS_MEPA : EVGDQHVGNISIGQGCAYKALIEHEILHALGFYHEQSRTRDRDDYVNIWWDQILSGYQHNFDTYDDSLIT : 135
HV_HMP2 : YLGRVGGILQNISLDDGCWGKGTIVHEIIGHAMGFHEQNRPRDRDOYITIRWENIPESKKHNFRLYSNSLVD : 140

*           160          *           180          *           200
HS_MEPA : DLNTPYDYESLMHYQPFSEFNKNASVPTITAKIPEFNSIIGQRLDFSAIDLRLNRMYNCTT : 196
HV_HMP2 : SLNSPYDYRSYMQYSKTAEGINDSV-TLDPKLGIF-QLGQRVGFTHEHQYQAMQLYRCQG : 199
    
```

C Astacin/Cnidaria

```

*           20           *           40           *           60           *
AA_AST  : AAILG--DEYLWSSGGVPIPYTFAG-VSGADQSAIILSGMQELEEKTCIRFVPRIT-----ESDYVEIFTSGS : 62
HEA2    : NAFASIKGGR-WPGAKIPYVVEGSISSRGISAINAAIADYHKFTCLRFRHRTN-----ERAHISFYK-GG : 63
PC_PMP1 : NAFASIKGGR-WPGAKIAMQIERSIGSGGRTAIAAIAQYHKHTCLRFRHKRTN-----ERSHISFYK-GG : 63
HV_FARM1 : -----MSSSNHIVLRAIDYEHKHTCLKFKVTRN-----QDAYLSFYF-GG : 40
HV_HMP1 : NQFKGVTLDRIWPGAQIPYVFDNDFNRRRQLVNOAITSYNQHTCVRFAPRTN-----QRDFVHMVN-EQ : 64
HEA4    : GLTKKSANLR-WPNNEMPYILSDNFSEDEKTSIRDAIDIIRSVTCIQFTEYPLTSP-PQ-DHVAIFP-GS : 66
HEA3    : LAPFDAIRGGVWPGGVVPYTFAGNFYK--AVWDAIKKEYAETCIRFKPYTASLAKKAGGYVEFMQ-GG : 67
HEA1    : AVVGVAMSQRLWPNGNVVPYVSTIKAANAINALK-AIEAN--TKCITETQKTAAD----KNFIEFKH-GG : 62

80           *           100          *           120          *           140
AA_AST  : GCWSYVGRISGAQQVSLQANGCVYHGTIIEHELMHATGFYHEHTRMDRDNVVTINYQNVDPSSMTSNFDDIDT : 132
HEA2    : GCSSPVGYRSGRMNRISLAGGWNKGIVMHEIGHSIGIYHEQSRPDRDNVVRILWNNISSRAKFNFNKQR : 133
PC_PMP1 : GCSSPVGYRSGRQNRISLAGGWRGTGIVMHEIGHSIGIYHEQSRPDRDSYVEIVWGNIPSSKKNFNKKS : 133
HV_FARM1 : GCSSLVGYVRCRINDVSLAGGCLRLGIVMHEIGHSIGIYHEQSRPDRDDHVTIIVWNNIQSNMRFNFDKFD : 110
HV_HMP1 : GCWSIIG-RQNSQKLLGDWCLDIGTAIHEMMHAAGFYHEQSRNDRDDHVVIFWNNIQRGEDYNFQKYR : 133
HEA4    : GCWSYVGRIGGAQQISLEHGGCINSGTIIEVLAHALGVHHEQARADRDGYVTINLLNIIIGREINFDKQD : 136
HEA3    : GCSSYIGRQRGRQQ-ISLANGGCTKGVAIHEMMHALGFYHEQSRDRDKYITIHWNINIKRMWFNFQKYR : 136
HEA1    : GCYSYVGMRRRTNPIYLSPPCYRSKDIVMHEVGHALGLNQHPG----- : 105

*           160          *           180          *           200          *
AA_AST  : Y--SRYGVEDYQYYSIMHYGKYSFISIQGVLETTIVPLQNGIDLTDYPDKAHMLQTDANQIINNYTNECS- : 199
HEA2    : TSNINSLGTPYDFRSMHYSFGAFGG---GRRPDEVK-DSRTPNLMGQRSGFSAIDIKQITIMY---CGG : 196
PC_PMP1 : SSSINSLGTPYDFRSMHYSSTAIGG---GRTTIRTK-NPSNQRLGQRSGFSEIDIKQINLMY---CNG : 196
HV_FARM1 : RNKINSLGFPDYESMMHYESNAFSG---GOVTRTK-DPSKQKLGNRQGFSEIDKQIINANYN---CNR : 174
HV_HMP1 : ---TEYYQTYDHNSIMHYRNDDESA--NGKNTIQAK--NNPG--LKLGNQFSKSDIAAIIQNYK---CSG : 193
HEA4    : ---TDSLGEVYDPKSVMHYGNDAFSN--GNGPTITWK--ADPT--LSLGGNELTEKDIEQMLFYE---CNA : 196
HEA3    : SGAASTLGEFYDKSSIMHYGSYAFSI--NRRKTIIVSK--SSPNERL-GQROGLSRIDKNQLNKYYS---CKK : 200
HEA1    : -----QKPSAP-----LG-KADFEKLNKMYT--C-- : 127
    
```

Fig. 2. Multiple alignments of the active protease domain of astacins, including all cnidarian astacins and selected examples from other organisms. Results were corrected with help of the known secondary structures and conserved regions of crayfish astacin. For abbreviations and Swissprot/TREMBL/PIR Accession nos. see Fig. 3.

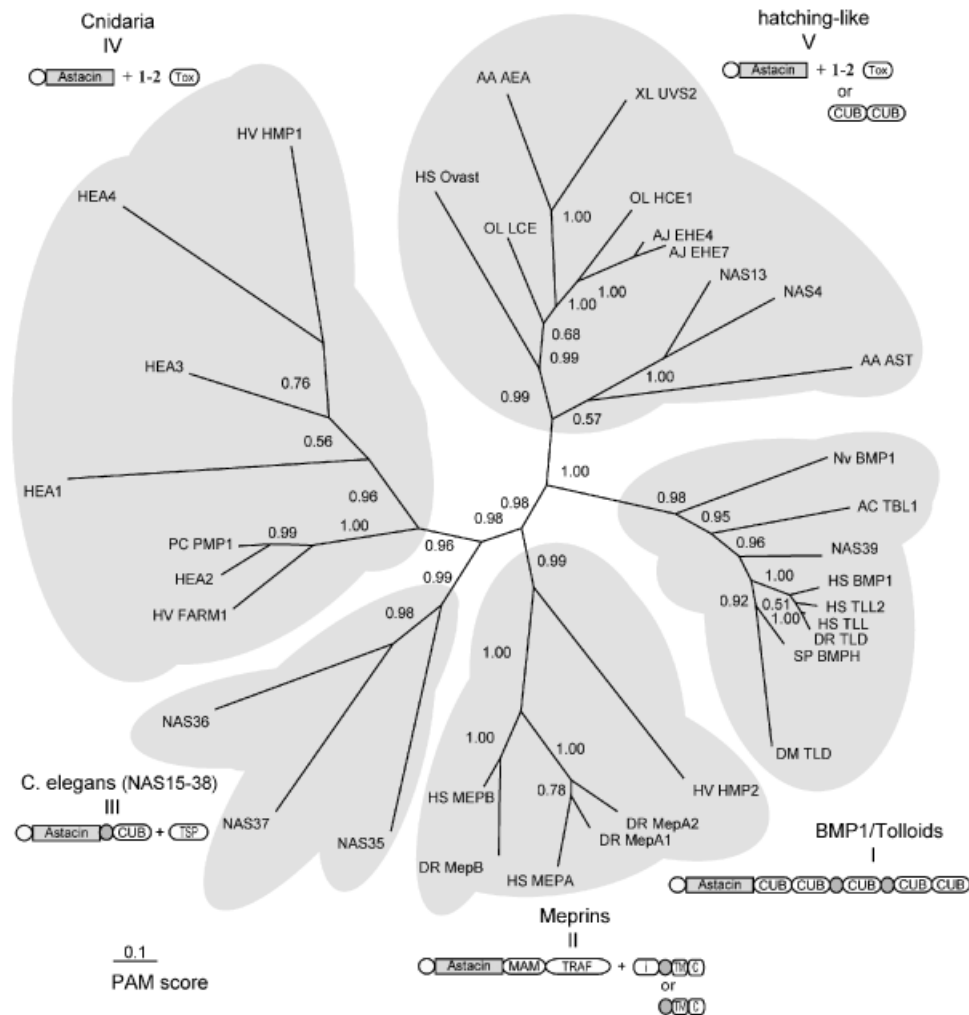


Fig. 3. Phylogenetic relationships of astacins in the Eumetazoa. The tree was deduced by Bayesian analysis based on the alignment of the amino-acid sequences of the catalytic chain, covering the region from Ala-1 to Leu-200 in the prototype, crayfish astacin. At branching points Bayesian posterior probabilities are given as an indication for the confidence of the tree presented. The main clusters are shaded: Cluster I, bone morphogenetic protein (BMP)/tolloids; cluster II, meprins; cluster III, *Caenorhabditis elegans* astacins; cluster IV, cnidarian genes; cluster V, hatching enzymes and their relatives. Typical domain structure of each cluster is depicted: ○, Pre-Pro-Sequences; Astacin active protease domains, covering the region from Ala-1 to Leu-200 in the prototype crayfish astacin (Titani et al. 1987, Geier et al. 1997), ●, EGF (epidermal growth factor)—like modules (PFAM accession number: PF00008); CUB, CUB modules, named after their occurrence in complement components C1r/C1s, embryonic sea urchin protein Uegf, and BMP-1 (Yamaguchi et al. 1995); TSP-1, Thrombospondin type 1-like domains (SMART accession number: SM0209); Tox ShK toxin domain (SMART accession number: SM0254); MAM, MAM domain of the meprins and *Hydra* HMP-2, named according to their presence in meprin, A5 receptor protein, tyrosine phosphatase μ and other proteins (SMART accession number: SM0137); other Meprin domains: TRAF, I-domain, intervening sequence, transmembrane domain, cytoplasmatic domain (for details, see Wolz and Bond 1995). The scale bar represents a distance of 0.1 accepted point mutations per site (PAM). Abbreviations and Swissprot/EMBL/PIR Acc. Nos. of the astacins: AA Ast, *Astacus astacus* astacin (P07584); AC TBL-1, *Aplysia californica* TBL-1 (P91972); AJ EHE4 and 7, *Anguilla japonica* EHE-4 (Q90Y89) and EHE-7 (Q90Y87); DM TLD, *Drosophila melanogaster* Tolloid (P25723); DR MepA1, *Danio rerio* novel protein similar to vertebrate meprin a (CAI21133); DR MepA2, *Danio rerio* novel protein similar to vertebrate meprin a (CAI21134); DR MepB, *Danio rerio* novel protein similar to vertebrate meprin b (CAI11689); DR TLD, *Danio rerio* Tolloid (AAC60304); HS BMP-1, *Homo sapiens* bone morphogenetic protein 1 (Q14874); HS MEP A and B, *Homo sapiens* Meprin a (Q16819) and b (Q16820); HS Ovast, *Homo sapiens* oocyte astacin (Ovastacin) AJ537600; HS TLL and TLL-2, *Homo sapiens* Tolloid like 1 (Q9NQS4) and 2 (Q9UQ00); HV FARM1, *Hydra vulgaris* astacin family metalloendopeptidase FARM-1 (Q9U4X9); HV HMP1, *Hydra vulgaris* Metalloprotease 1 (U22380); HV HMP-2, *Hydra vulgaris* Metalloprotease 2 (Q9XZG0); NAS4, 13, 35, 36, 37, *C. elegans* nematode astacin 4, 13, 35, 36, 37 (see www.wormbase.org); OL LCE and HCE1, *Oryzias latipes* low choreolytic enzyme (P31581) and high choreolytic enzyme 1 (EMBL: M96170); PC PMP-1, *Podocoryne carnea* Metalloprotease 1 (O62558PL); SP BMPH, *Strongylocentrotus purpuratus* BMP-1 homolog (P98069); XL UVS2, *Xenopus laevis* embryonic protein UVS.2 (P42664).

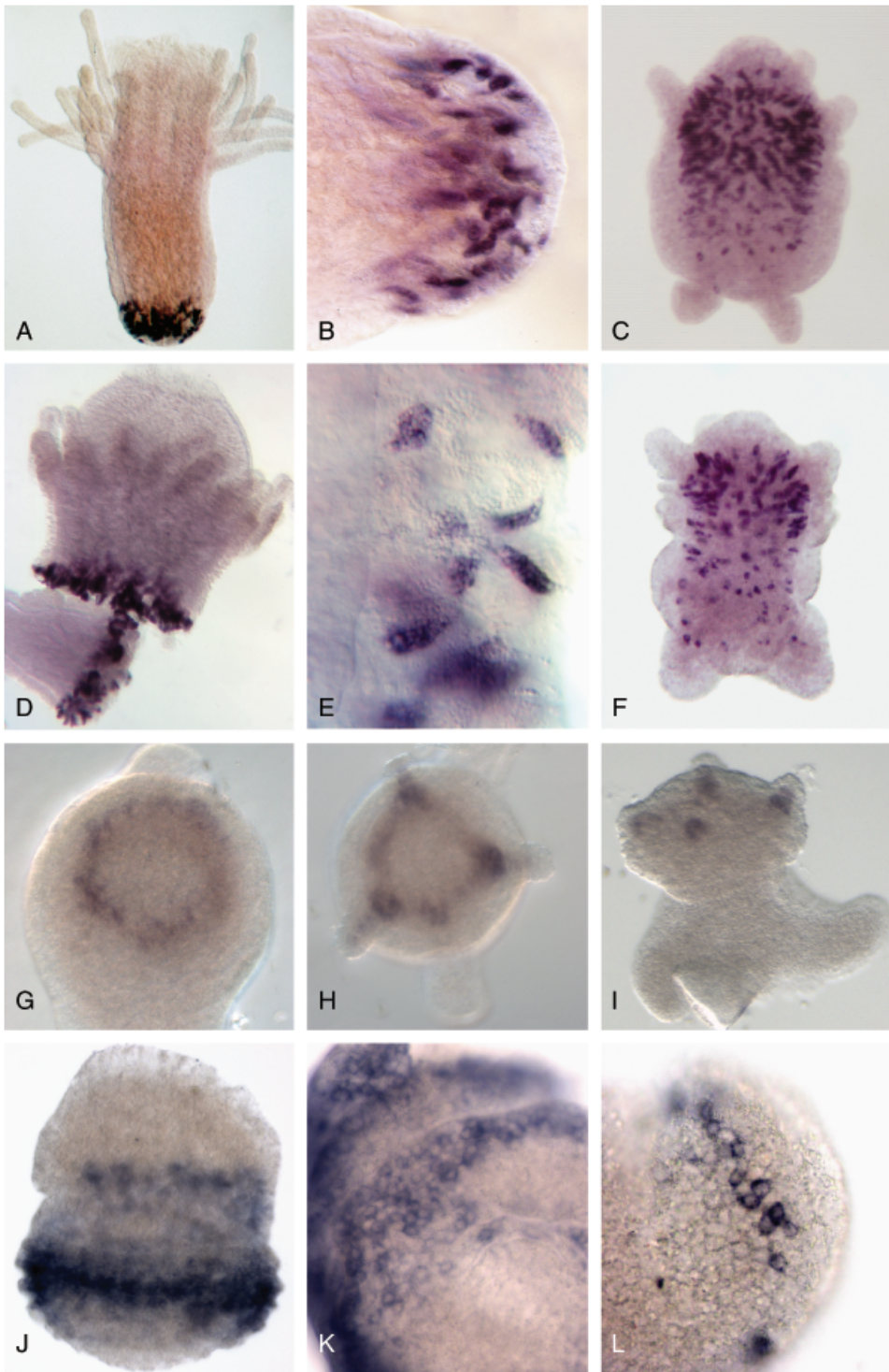


Fig. 4. In situ hybridization of the 4 *Hydractinia astacins* during metamorphosis. (A–C) *Hydractinia echinata* astacins (*HEA1*), (D–F) *HEA2*, (G–I) *HEA3*, (J–L) *HEA4*. (A) Injured polyp, 2 h following removal from the colony showing strong *HEA1* expression at the injury site. (B) A higher magnification of the injury area. (C) Metamorphosing polyp, 18 h post CsCl induction. (D) Upper and lower parts of a polyp, 2 h following injury. *HEA2* is expressed at the injury area. (E) Higher magnification of the injury area. Gland cells expressing *HEA2* are visible near morphologically similar cells not expressing the gene. (F) Metamorphosing polyp, 18 h post CsCl induction. (G) Early metamorphosing polyp, about 10 h post-CsCl treatment. *HEA3* expressing cells form a ring around the future mouth. (H) Several hours later, staining concentrates in tentacle buds. (I) Metamorphosing polyp, 14 h following CsCl treatment. *HEA3* is expressed in the endoderm of tentacle buds. (J) Early metamorphosing specimen, about 8 h following CsCl treatment. *HEA4* expressing cells form two bands around the body column. (K) A higher magnification of the previous image looking from above. (L) Higher magnification of interstitial cells expressing *HEA4*.

solely in a fraction of the endodermal gland cells, mostly at the upper body column of the metamorphosing polyp, starting just below the hypostome, decreasing in a gradient toward the future stolonal compartment (Fig. 4). In addition to the developmentally regulated expression, *HEA1* was also identi-

fied as an early-responding gene following injury (Fig. 4). *HEA1*'s transcripts were already detectable about 30 min following the infliction of tissue lesions of all types and at all positions in the polyp. Following injury, the gene maintained local high expression levels for about 24 h, after which it was

gradually silenced as the wound healed. As during metamorphosis, the expressing cells were only endodermal gland cells of a particular lineage, indistinguishable morphologically from other, adjacent gland cells, not expressing the gene (Fig. 4).

HEA2 showed an expression pattern very similar to that of *HEA1* (Fig. 4). Although we did not perform double labeling of the two genes, the in situ hybridization results performed using *HEA1* and *HEA2* cRNA probes were indistinguishable in spatial and temporal expression pattern. This held true not only for the ontogenetic component of expression but also for the role in the wound healing process where *HEA2* was expressed in a similar time frame following injury as was *HEA1*. Finally, both genes were expressed by the same cell type.

HEA3 was expressed during metamorphosis in endodermal cells around the budding tentacles (Fig. 4). Initially, about 10 h post-CsCl treatment, *HEA3* expressing cells formed a ring around the future mouth at the predicted tentacle budding zone (Fig. 4). About 4–6 h later, when tentacle buds became more evident, the ring-like expression pattern was replaced by a strong staining of the endoderm in the tentacle buds themselves, while the staining at the inter-tentacle tissue faded out (Fig. 4). Unlike *HEA1* and *HEA2*, *HEA3* was not up-regulated following injury. *HEA3* was also expressed occasionally in sexual polyps at the tentacle rudiments or in the head (not shown).

HEA4 was also most highly expressed during metamorphosis, similar to the other three *Hydractinia* astacins. In contrast to *HEA1*, 2, and 3, however, transcripts of *HEA4* were only detectable in the ectodermal layer. Moreover, *HEA4* was also strongly expressed in post-metamorphic colonies. During early metamorphosis, *HEA4* cRNA probes stained two bands around the metamorphosing polyp (Fig. 4). One band was at the position of the future tentacle crown zone, while the other demarcated the position of the future stolons. In post-metamorphic colonies, *HEA4* was most prominently expressed in stolon tips (not shown). The cells expressing *HEA4* were a sub-population of the interstitial cell lineage, possibly nematoblasts—the stinging cells' precursors—according to their position. This cell type was the only one observed to express the gene in both metamorphosis and in the post-metamorphic growth phase. Similar to *HEA3*, *HEA4* was not expressed during wound healing, in contrast to *HEA1* and *HEA2*.

Incubation with sense RNA probes, at the same RNA concentration and under the same conditions used for anti-sense hybridization, did not reveal any signal in any of the four genes.

DISCUSSION

We have cloned and characterized four members of the astacin family of zinc endopeptidases from the colonial marine

hydroid *H. echinata*. We have also performed a database search in the draft genome sequence of the sea anemone *N. vectensis*, a search revealing several astacin-like sequences including a BMP-1 homolog, a putative member of the meprin group, and some other astacins. Only the BMP-1-like sequence was of a quality enabling a reliable prediction of the entire processed protein. We then performed a phylogenetic analysis with astacins representing the main metazoan clades and astacin subgroups.

The predicted primary structure of the *Hydractinia* astacins showed secreted proteins with well-conserved astacin-like catalytic domains. *HEA1* formed an exception to this rule in having an incomplete active center in its catalytic domain. This protein, missing two of the zinc-binding amino-acid residues (Fig. 2), is expected to lack protease activity according to the current structural and biochemical understanding of astacins (Bode et al. 1992; Gomis-Rüth et al. 1993; Stöcker et al. 1993; Yiallourous et al. 2000). *HEA1* was also the only *Hydractinia* astacin lacking any additional protein domain. The other *Hydractinia* astacins all contained at least one ShKT domain (two such domains in *HEA3*). The ShK toxin domain, (SMART accession number: SM0254) was originally found in metridin, a toxin from a sea anemone, and was also called six-cysteine (SXC) domain. The function of this protein domain, however, is still unclear. Interestingly, it is also found in *Hydra* astacins and in some *C. elegans* astacins (Möhrlen et al. 2003). *HEA3* formed an exception in a different aspect: it had no detectable signal peptide according to the search algorithm of the Signal IP server. Nevertheless, one cannot exclude that this protein is secreted via an alternative secretory pathway.

Our phylogenetic analysis involved members of the astacin-like proteins representing the Eumetazoa from different organizational levels up to mammals. We included all known cnidarian astacins and six *C. elegans* proteins, as this animal is unusual, having a genome encoding 40 astacin genes (Möhrlen et al. 2003). The analysis revealed five clusters within the Eumetazoa: (1) the BMP-1/tolloids, (2) meprins, (3) *C. elegans*-specific astacins, (4) Cnidaria-specific astacins, and (5) the hatching enzyme-like astacins. The clusters obtained by the alignment of only the catalytic domains encompassed proteins with similar C-terminal domain structure and putative function, similar to a previous study, done with astacins of only higher animals (Möhrlen et al. 2003). The analysis included eight characterized and one predicted cnidarian astacins. Seven of the characterized cnidarian proteins clustered together as a Cnidaria-specific group (cluster #4). The *Hydra* protein HMP-2 was one exception, clustering with the meprins. Interestingly, although this analysis was based on the sequence of only the catalytic domain, HMP-2 also resembles other meprins in having a MAM domain, absent in all other astacins. Meprins have been identified so far only in vertebrates. The genomes of the two prominent invertebrate model

organisms, *Drosophila* and *C. elegans* do not encode meprins (Möhrlen et al. 2004). The lack of meprins in flies and worms may be because of gene loss in either the ecdysozoan ancestor, or independently in several lineages. Extensive gene loss in these taxa has already been documented previously (Guduric-Fuchs et al. 2004; Kortschak et al. 2004). The predicted *Nematostella* BMP-1/tolloid homologue was the second exception. Containing all typical domains found in other BMP-1/tolloid homologs, this protein clustered accordingly in this group in our phylogenetic tree and not with other cnidarian astacins. Again, clustering was based on the catalytic domain alone. The evolution of cnidarian members of these groups was, thus, typical of astacins, that is, similarity in the catalytic domain reflected similarity in the C-terminal domain structure. This mode of evolutionary conservation within a given protein family is striking. As a comparison, phylogenetic analysis of the TNF receptor-associated factor (TRAF) family, based on the TRAF domain alone, clusters together genes of different overall domain architectures (Mali and Frank 2004).

In situ hybridization of the four *Hydractinia* astacins suggested a developmental function to all, as their main expression pattern was developmentally regulated. *HEA1* and *HEA2* also showed a role in wound healing, a function recently assigned also to the closely-related matrix metalloproteases of mammals (Toy 2005). *HEA1* and *HEA2* had the same spatial and temporal expression pattern in development and in wound healing. They were also expressed by the same cell type. The sequence of the catalytic domain of *HEA1* seemed to be truncated, and its gene product is therefore expected not to act as a protease. This function may be fulfilled by *HEA2*, which clustered most closely to it within the *Hydractinia* genes (Fig. 3). It is unclear how ancient this mutational deletion is. Studying astacins in the closely related hydroid, *Podocoryne*, a sister genus to *Hydractinia*, in which only one astacin—almost identical to *HEA2* (Fig. 2)—has been identified so far, may shed light on this point. The function of *HEA2* could be in the processing of the mesoglea, the cnidarian basal lamina, during growth of the metamorphosing polyp. The function during wound healing may be similar; the mesoglea has to be “rearranged” in the wound area. *HEA3* was expressed in budding tentacles. However, the apparent lack of a signal peptide in this protein makes a function similar to *HEA1* and 2, that is the processing of the mesoglea in pattern formation, questionable. Nevertheless, the restricted, ontogenetically regulated expression pattern of *HEA3* indicates that it does fulfill a role in development. *HEA4* was the only gene expressed in the ectoderm and also displayed a prominent expression in post-metamorphic life. It was expressed by a subpopulation of interstitial stem cells. In hydroids, interstitial cells, or i-cells, give rise to stinging cells (nematocytes), nerve cells, gland cells and germ cells (Bode 1996). In *Hydractinia*, i-cells are totipotent and may also dif-

ferentiate into epithelial cells (Müller et al. 2004). *HEA4* expressing cells may be nematoblasts, the stinging cells’ precursors, according to their position during metamorphosis. All i-cells, however, are migratory cells. We propose that *HEA4* functions to facilitate their movement in the interstitial spaces of the ectoderm.

The other four characterized cnidarian astacins also act during development, although they seem to fulfill a wide variety of tasks. The predicted *Nematostella* BMP-1/tolloid homologue has not been characterized yet. However, its strong structural conservation with other BMP-1/tolloid-like proteins suggests that it also functions during embryonic development or matrix differentiation. We suggest that BMP-1/tolloids and meprins, which are remarkably conserved throughout the Eumetazoa, have diverged early in metazoan evolution to generate two groups of homologous proteins with distinct structure and functions. The hatching enzymes, characterized by a simple C-terminal structure, have probably evolved later, after the divergence of the Cnidaria. This group is more diverse than the BMP-1/tolloids and the meprin groups. Its members have broad substrate specificity, in contrast to the specific substrates of BMP-1/tolloids, and mostly function in the breakdown of the egg envelope during hatching, or in food digestion. Clusters #4 and #5 included only cnidarian and *C. elegans* genes, respectively. The developmental function assigned to the characterized cnidarian astacins seems therefore to be phylum-specific, phylogenetically unrelated to the morphogenetic function of BMP-1. These genes, and the *C. elegans*-specific astacins, are probably the outcome of several gene duplication events, which occurred after their divergence from the main line of metazoan evolution.

Acknowledgments

Thanks are due to Gabi Günther, Susanne Kremer, Jan Drexel, and Nicole Ungerer for technical assistance. Hans Bode (University of California, Irvine) has been collaborating in the *Hydractinia* EST project. Joseph Ryan and Jim Mullikin (Boston University) assisted in the search for astacins in the draft genome sequence of *N. vectensis*. Werner A. Müller (University of Heidelberg) is kindly acknowledged for reviewing an early version of the manuscript and for drawing Fig. 1. This study was supported by grants from the DFG, the German Research Council.

REFERENCES

- Bankus, J. M., and Bond, J. S. 1996. Expression and distribution of meprin protease subunits in mouse intestine. *Arch. Biochem. Biophys.* 331: 87–94.
- Barrett, A. J., Rawlings, N. D., and Woessner, J. F. 2004. *Handbook of Proteolytic Enzymes*. Academic Press, London.
- Blader, P., Rastegar, S., Fischer, N., and Strahle, U. 1997. Cleavage of the BMP-4 antagonist chordin by zebrafish tolloid. *Science* 278: 1937–1940.
- Bode, H. R. 1996. The interstitial cell lineage of hydra: a stem cell system that arose early in evolution. *J. Cell. Sci.* 109: 1155–1164.

- Bode, W., Gomis-Rath, F., 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.
- Bond, J. S., and Beynon, R. J. 1995. The astacin family of metalloendopeptidases. *Protein Sci.* 4: 1247–1261.
- Dumermuth, E., et al. 1991. The astacin family of metalloendopeptidases. *J. Biol. Chem.* 266: 21381–21385.
- Fan, T. J., and Katagiri, C. 2001. Properties of the hatching enzyme from *Xenopus laevis*. *Eur. J. Biochem.* 268: 4892–4898.
- Frank, U., Leitz, T., and Müller, W. A. 2001. The hydroid *Hydractinia*: a versatile, informative cnidarian representative. *BioEssays* 23: 963–971.
- Gajewsky, M., Leitz, T., Schlossherr, J., and Plickert, G. 1996. LWamides from Cnidaria constitute a novel family of neuropeptide with morphogenetic activity. *Roux's Arch. Dev. Biol.* 205: 232–242.
- Geier, G., Jacob, E., Stöcker, W., Zwilling, R. 1997. Genomic organization of the zinc-endopeptidase astacin. *Arch. Biochem. Biophys.* 337: 300–307.
- 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., 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.
- Guduric-Fuchs, J., Möhrlen, F., Frohme, M., and Frank, U. 2004. A fragile X mental retardation-like gene in a cnidarian. *Gene* 343: 231–238.
- Hishida, R., Ishihara, T., Kondo, K., and Katsura, I. 1996. hch-1, a gene required for normal hatching and normal migration of a neuroblast in *C. elegans*, encodes a protein related to TOLLOID and BMP-1. *EMBO J.* 15: 4111–4122.
- Huelsenbeck, J. P., and Ronquist, F. 2001. MRBAYES: bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
- Köhler, D., Kruse, M., Stöcker, W., and Sterchi, E. E. 2000. Heterologously overexpressed, affinity-purified human meprin alpha is functionally active and cleaves components of the basement membrane in vitro. *FEBS Lett.* 465: 2–7.
- Kortschak, R. D., Samuel, G., Saint, R., and Miller, D. J. 2004. EST analysis of the cnidarian *Acropora millepora* reveals extensive gene loss and rapid sequence divergence in the model invertebrates. *Curr. Biol.* 13: 2190–2195.
- Kumpfmüller, G., Rybakine, V., Takahashi, T., Fugisawa, T., and Bosch, T. C. G. 1999. Identification of an astacin matrix metalloprotease as target gene for *Hydra* foot activator peptides. *Dev. Genes Evol.* 209: 601–607.
- Mali, B., and Frank, U. 2004. Hydroid TNF-receptor-associated factor (TRAF) and its splice variant: a role in development. *Mol. Immunol.* 41: 377–384.
- Möhrlen, F., Bond, J. S., and Stöcker, W. 2004. Other astacin homologues. In, A. J. Barrett, N. D. Rawlings, and J. D. Woessner (eds.). *Handbook of Proteolytic Enzymes*. Academic Press, San Diego, CA.
- Möhrlen, F., Hutter, H., and Zwilling, R. 2003. The astacin protein family in *Caenorhabditis elegans*. *Eur. J. Biochem.* 270: 4909–4920.
- Müller, W. A., and Leitz, T. 2002. Metamorphosis in the Cnidaria. *Can. J. Zool.* 80: 1755–1771.
- Müller, W. A., Teo, R., and Frank, U. 2004. Totipotent migratory stem cells in a hydroid. *Dev. Biol.* 275: 215–224.
- Mullins, M. C. 1998. Holy tolloido: tolloid cleaves SOG/chordin to free DPP/BMPs. *Trends Genet.* 14: 127–129.
- 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.
- Reddi, A. H. 1996. BMP-1: resurrection as procollagen C-proteinase. *Science* 271: 463.
- Shibata, 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: 8349–8354.
- Stöcker, W., Gomis-Rüth, F., 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., et al. 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.
- Suzuki, N., et al. 1996. Failure of ventral body wall closure in mouse embryos lacking a procollagen C-proteinase encoded by *Bmp1*, a mammalian gene related to *Drosophila* tolloid. *Development* 122: 3587–3595.
- Titani, K., Torff, H. J., Hormel, S., Kumar, S., Walsh, K. A., Rodl, J., Neurath, H., and Zwilling, R. 1987. Amino acid sequence of a unique protease from the crayfish *Astacus fluviatilis*. *Biochemistry.* 26: 222–226.
- Toy, L. W. 2005. Matrix metalloproteinases: their function in tissue repair. *J. Wound Care.* 14: 20–22.
- Wolz, R. L., and Bond, J. S. 1995. Meprins A and B. *Methods Enzymol.* 248: 325–345.
- Yamaguchi, T., Fukase, M., Sugimoto, T., and Chihara, K. 1995. Hydrolysis of a carboxy-terminal fragment of parathyroid hormone-related protein by rat kidney: evidence for a crucial role of meprin. *Horm. Metab. Res.* 27: 131–136.
- Yan, L., Fei, K., Zhang, J., Dexter, S., and Sarras, M. 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. J. 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., et al. 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.
- Yiallourou, I., Grosse-Berkhoff, E., and Stöcker, W. 2000. The roles of Glu93 and Tyr149 in astacin-like zinc peptidases. *FEBS Lett.* 484: 224–228.
- Zwilling, R., and Stöcker, W. 1997. *The Astacins: Structure and Function of a New Protein Family*. Verlag Dr. Kovac, Hamburg.