ISSN 0006-2979, Biochemistry (Moscow), 2012, Vol. 77, No. 2, pp. 119-127. © Pleiades Publishing, Ltd., 2012.
Original Russian Text © N. P. Balaban, N. L. Rudakova, M. R. Sharipova, 2012, published in Biokhimiya, 2012, Vol. 77, No. 2, pp. 149-159.

REVIEW

Structural and Functional Characteristics and Properties of Metzincins

N. P. Balaban, N. L. Rudakova*, and M. R. Sharipova

Kazan (Volga Region) Federal University, ul. Kremlevskaya 18, 420008 Kazan, Russia;
fax: (843) 292-4448; E-mail: public.mail@ksu.ru; natalialrudakova@gmail.com

Received January 23, 2011
Revision received June 21, 2011

Abstract—In this review the main families of endopeptidases belonging to the clan of metzincins of zinc-dependent metalloproteinases in organisms of wide evolutional range from bacteria to mammals are considered. The data on classification, physicochemical properties, substrate specificity, and structural features of this group of enzymes are given. The activation mechanisms of metzincins, the role of these proteins in organisms, and their participation in various physiological processes are discussed.

DOI: 10.1134/S0006297912020010

Key words: metalloproteinases, metzincins, astacins, adamalysins/reprolysins, matrixins, serralysins

Proteolytic enzymes are represented in genomes of organisms over a wide evolutional range from procaryotes to higher eucaryotes. Being strategic cell enzymes, they play a key role in vital activity of organisms. This is evidenced by the fact that more than 640 genes encoding peptidases or their homologs are identified in the human genome [1]. These enzymes function in cells not only as degradation enzymes, but they also participate in regulatory processes, and violation of such processes results in initiation and development of pathology [2]. Studies of the physicochemical and catalytic properties of these enzymes and their structural and functional roles *in vivo* allow classification of these proteins and determine their practical importance and perspectives of their use.

Proteolytic enzymes (EC 3.4) are commonly divided into six classes by their catalytic mechanism: aspartate, glutamate, metalloproteinase, cysteine, serine, threonine, and metalloproteinases (endopeptidases, metalloendopeptidases) being the most numerous and varied classes among them (EC 3.4.24) [3, 4]. According to this classification, zinc-dependent metalloproteinases with the active site motif HEXXH are distributed over families from M1 to M13 dependent on the structural properties of protein molecule and its catalytic and enzymatic properties and are pooled to form the MA clan. Enzymes with different active site motif belong to other families [5, 6]. Many zinc-dependent metalloendopeptidases contain calcium ions that play an important role in stabilization of the spatial structure of the enzyme. These enzymes are found in animals including humans, and they are produced by bacteria, Actinomycetes, and microscopic fungi.

In 2003, zinc-dependent metallopeptidases were divided into two large groups: zincins with the conservative active site sequence **H**EXX**H** with two histidines as zinc ligands and glutamate as a catalytic base, and inverzincins having a reversed sequence of amino acid residues of the active site **H**XXE**H** (Fig. 1) [2, 7]. The suggested scheme of division of proteinases dependent on position and context of zinc ligands is in accord with the common classification presented in the peptidase database MEROPS (http://merops.sanger.ac.uk).

In turn, zincins are divided into three clans dependent on their nature and position of the third zinc ligand: gluzincins, aspzincins, and metzincins (Fig. 1) [7]. Thermolysin having the second consensus motif N**E**XXSD located near the C-end of the **HE**XX**H** motif of the active site and containing a glutamate residue as the third zinc ligand is a typical representative of gluzincins. The structure of the active site of enzymes belonging to clan of aspzincins is similar to that of gluzincins: the glutamate residue of the second motif is replaced by aspartate. This raised the suggestion that aspzincins can form a sub-clan of gluzincins. The clan of metzincins containing an extended motif of the active site **HEXXHXXGXXH/D**

* To whom correspondence should be addressed.

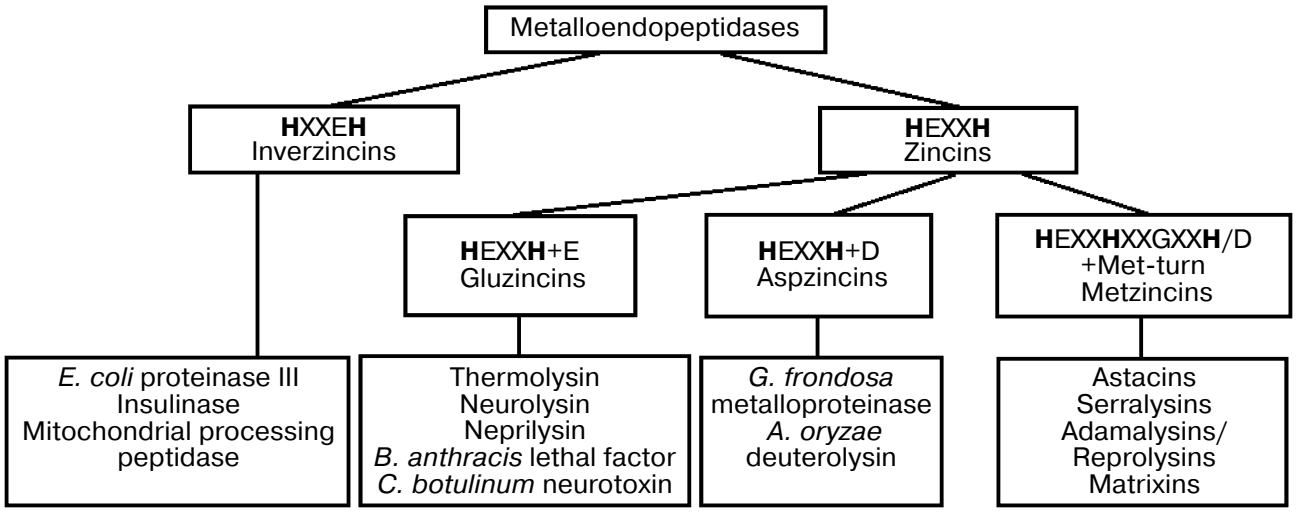

Fig. 1. Classification scheme of metalloendopeptidases based on active site structure.

with three histidines as zinc ligands is the most representative among zincins [2, 7]. The presence of a Met-turn with a conservative methionine located in a 1,4-β-loop in molecular structure is a unique feature of this clan; the Met-turn is used for correct classification of this group of enzymes. The corresponding endopeptidases are named metzincins due to the presence of the conservative methionine residue in this sequence.

CHARACTERISTICS OF ENZYMES BELONGING TO CLAN OF METZINCINS

The clan of metzincins (MEROPS clan MA(M)) includes the following families of endopeptidases: astacins, serralysins, matrixins, adamalysins/reprolysins, snapalysins, leishmanolysins, and others with common structural features mentioned above (Fig. 1). Physicochemical properties of these enzymes, their substrate specificity, structural features of the catalytic domain, and multi-domain structure of proteins are well studied. Along with the above-mentioned families, endopeptidases with unknown spatial structure by now have been discovered: fragilysins (MEROPS M10C), pappalysins (MEROPS M46), gametolysins (MEROPS M11), turingelysins (MEROPS M6), archaemetzincins, coelysin, helicolysins, and others [8]. The biochemical properties, substrate specificity, substrates, and inhibitors of these proteinases are quite well studied. These enzymes also contain an extended zinc-binding motif of the active site with conservative amino acid residues typical of metzincins and Met-turn sequence with invariant methionine. Based on this, they are assigned to the clan of metzincins and included in the MEROPS peptidase database.

Family of astacins. Astacin isolated in 1967 from the alimentary track of the crayfish *Astacus astacus* L. was the first studied enzyme of the astacin family; these enzymes are mainly found in representatives of the animal kingdom [9]. By now these enzymes have not been found in plants and fungi, but one astacin-like enzyme (flavastacin) was found in gram-negative bacteria [10]. Secreted and membrane-bound meprins α and β of humans and rodents, mice procollagen C-endopeptidase, human bone morphogenetic protein 1 (BMP1), choriolysins L and H of salmons, and other endopeptidases belong to the family of astacin-like endopeptidases [11-15].

Astacin-like endopeptidases are synthesized as inactive precursors with a signal peptide and propeptide at the N-end of molecule. On proteolytic removal of the N-terminal propeptide, the zymogen is activated via the "tyrosine switch" mechanism, which differs from the "cysteine switch" mechanism of matrixins and adamalysins [16].

The mature astacin molecule (MEROPS M12.001) is a protease domain consisting of 200 amino acid residues (a.a.) with molecular mass 20.3 kDa with a typical extended motif of the active site and conservative structure of Met-turn SX**M**HY [17]. Astacin is stable in neutral and weak acidic media, but at pH < 4.0 the protein is irreversibly inactivated. The pH optimum of the enzyme is 7.5-9.5 and depends to some extent on substrate. Casein, azocasein, gelatin, collagen, and B-chain of oxidized insulin are substrates for astacin; on their hydrolysis the enzyme gives preference to amino acid residues with short side chains (Ala, Thr, Ser, Gly) at P1′ position, to proline at P2 and P3 positions, and to hydrophobic residues at P3′ and P4′ positions. Nitroamide-like substrates 5-7 a.a. in length and fluorescent substrates can be synthetic substrates for astacin-like

endopeptidases. Astacin and astacin-like endopeptidases are insensitive to inhibitors of serine, cysteine, and aspartate proteinases, phosphoramidon (an inhibitor of thermolysin-like proteinases) and tissue inhibitors of metalloproteases (TIMPs). The activity of these enzymes is inhibited by 1,10-phenanthroline and EDTA, but the inhibitory effect of 1-10 mM EDTA manifests itself only after extended incubation [18, 19].

Beside astacin, the structure and properties of many astacin-like proteinases (meprins, BMP1, TLD-tolloid endopeptidase from *Drosophila melanogaster*, and others) have been studied. Unlike the simple structure of crayfish astacin, these proteins have multi-domain structure, but the protease domain has a structure similar to that of astacin. Additional non-protease domains (CUB-domain, EGF-domain, and others) participate in multimerization but are not necessary for catalysis and secretion. They promote cell adhesion and provide protein–protein interaction and/or binding to calcium ions, forming rigid structures in certain portions of the protein molecule, and this affects the structural and functional properties of these enzymes [7, 18, 19].

Flavastacin (MEROPS M12.066), an astacin-like endopeptidase, was isolated from culture liquid of the gram-negative bacterium *Flavobacterium meningosepticum* in 1995 [10]. The flavastacin gene was cloned and sequenced, and its expression was studied [20]. The molecular mass of flavastacin is 38.8 kDa, cysteine residues are absent from protein molecule, and the N-terminal propeptide sequence and signal peptide are at the N-end of the protease domain. The amino acid sequence of the protease domain is 20% identical to astacin and 24-29% identical to meprins. The protease domain of flavastacin has a highly conservative motif of the active site **H**EIM**H**SMGIM**H**E and a methionine turn SVMMY. The presence of a glutamate residue located after the third zinc ligand histidine and tyrosine residue in the Met-turn ascribe it to the family of astacin-like endopeptidases [19]. Discovery of microbial astacin-like endopeptidase suggests that a common ancestral gene of this protein existed before the beginning of divergence of this family of enzymes.

Some properties of homogeneous flavastacin preparations have been studied. This enzyme with pH optimum in the neutral range performs limited proteolysis of peptides via aspartic acid residues at P1′ position. This was proved by hydrolysis of a series of synthetic peptides of FA-Leu-Ala-Asp-Ala-Ser-NH_2 type. The same substrates with the Asp residue replaced by Glu were hydrolyzed only at high enzyme concentrations and on extended incubation.

Astacin-like endopeptidases *in vivo* participate in development processes including embryo differentiation and morphogenesis, remodeling, and differentiation of tissues, and they promote formation of chondral tissue and bones and also biosynthesis of collagen [7]. A large group of astacin-like proteinases (BMP1 proteinase – bone morphogenetic protein 1 and colloidal endopeptidases from *Drosophila melanogaster*) play a key role in formation and assembling of extracellular matrix [21]. A group of so-called hatching enzymes is responsible for degradation of embryonal film at hatching of crustaceans, fishes, frogs, birds, and other animals [2]. The function of meprins is related with development of organisms. The level of digestive meprin α in mice in the nursing period is higher than that of meprin β. The end of the nursing period and change for hard food is characterized by decrease in meprin α level and increase in meprin β level. The meprin level drastically decreases when such human diseases as hydronephrosis, nephropathia, and nephritis arise and develop and also on kidney transplantation. All these diseases and states result in death of tissue cells, and meprin level can thus serve as a diagnostic test [2].

The serralysin family (MEROPS M10B) includes such bacterial enzymes as serralysin *Serratia* (MEROPS M10.051), aeruginolysin *Pseudomonas* (MEROPS M10.056), mirabilysin *Proteus* (MEROPS M10.057), endopeptidases A, B, C *Erwinia* (MEROPS M10.052, M10.053 and M10.054, respectively), endopeptidase PrtA *Photorhabdus* (MEROPS M10.063), and other endopeptidases. The name of a particular enzyme in most cases is formed from the names of the bacterial species and enzyme, e.g. *Serratia* and *lysin*. The representatives of this family of proteins have similar physicochemical properties, structural organization, mechanism of protein secretion, and activation of the protein domain.

Serralysins are often called alkaline metalloproteinases because they are active over a wide pH range from 6.0 to 10.0 with pH optimum 8.0. Casein, azocasein, and gelatin and its derivatives are substrates for serralysins. Synthetic *p*-nitroanilides 6-8 residues in length also are good substrates for these enzymes. Serralysin and aeruginolysin hydrolyze these substrates with preference for Gly, Ala, and Arg residues and also for hydrophobic residues at P1′, P2, and P2′ positions [22]. Endopeptidase PrtA from the pathogenic insect *Photorhabdus luminescens* hydrolyze both synthetic hexapeptides and natural substrates via the bonds of Val, Ala, and Leu residues. The A- and B-chains of oxidized insulin, IgA and IgG, protective and cytoskeletal proteins, complement system components, and extracellular matrix proteins are natural substrates for this proteinase, the rate of hydrolysis of these substrates by endopeptidase PrtA being higher than by other serralysins [23].

Activity of serralysins is suppressed by EDTA, 1,10-phenanthroline, peptidyl mercapto anilides, and protein inhibitors (10-11 kDa) isolated from the same *Serratia*, *Pseudomonas*, and *Erwinia* strains [7, 24]. These enzymes are not inhibited by diisopropylfluorophosphate and phosphoramidon.

The molecular masses of enzymes of serralysin family are 49-55 kDa, p*I* 4.1-4.5, and the primary structures

of the protease domains are 50-55% identical to each other. The tertiary structures of serralysins *Serratia* sp. SM6 and E-15, aeruginolysin *Pseudomonas aeruginosa*, mirabilysin *Proteus mirabilis*, and proteinase *Erwinia chrysanthemi* are known [25-28]. The enzymes have an extended active site motif with three histidine ligands and conservative methionine turn SXMXY containing a Tyr residue. On activation of serralysins, tyrosine becomes the fifth zinc ligand and is included in substrate binding and/or stabilization of transition state during catalysis. Serralysins are secreted as auto-activating zymogens with a propeptide of 9-20 a.a. at the N-end of protease domain.

Structural domains of serralysins (N- and C-domains) each contain ~220 a.a. [7]. The N-domain is similar to the protease domain of astacin and requires correct folding. The amino acid sequence Gly-Gly-Xaa-Gly-Asn-Asp located in the C-terminal domain plays an important role in binding of Ca^{2+} ions. The C-domain is necessary for secretion of the enzyme, because serralysins are synthesized without a signal peptide [7]. The C-end of this domain contains from 20 to 50 a.a. acting as signals for translocation, in which the conservative C-terminal motif DXXX (X – any hydrophobic amino acid) is necessary for initiation of secretion. For serralysins and mirabilysins, this sequence is DFIV, which is not proteolytically removed on secretion of serralysins into the medium [26].

One function of serralysins is supposed to be related with virulence factor causing various infections of newborns and infections after surgical operations. These enzymes participate in initiation of such diseases as endocarditis, nephropyelitis, dermatitis, sepsis, and pneumonia, and their functioning results in degradation of soft tissues and membranes of digestive epithelial cells and also participate in initiation of diseases similar to those in bacterial infection: inflammatory processes, tetanus, botulism, gangrenous emphysema, bacterial meningitis, and anthrax [7, 8]. Possessing virulent potential, serralysins can act similar to plasminogen as an anticoagulative blood factor and thus can be used for treatment of thrombosis. However, these enzymes can inactivate protective proteins (e.g. immunoglobulins and γ-interferon), protein inhibitors, lysozyme, transferrin, and complement proteins, which results in exhaustion of host protective forces [29]. These properties of serralysins hinder development of drugs based on them.

The matrix metallopeptidase family (matrixins) (MEROPS M10A) includes secreted or membrane-bound endopeptidases found in higher mammals, fishes, amphibians, insects, plants, procaryotes, and viruses [8]. These enzymes are divided into three groups: true collagenases capable of limited proteolysis of the triple chain of collagen; gelatinases hydrolyzing collagen and gelatin; stromelysins with distinct proteolytic activity degrading proteoglycans [7]. Isolated from tissues of various animals, these endopeptidases have various names, so the general name matrix metalloproteinases (MMPs) came into use. Intra-tissue collagenase 1 (MMP-1, MEROPS M10.001), neutrophilic collagenase 2 (MMP-8, MEROPS M10.002), collagenases 3 (MMP-13, MEROPS M10.015) and 4 (MMP-18, MEROPS M10.018), gelatinases A (MEROPS M10.033) and B (MEROPS M10.034), stromelysins 1, 2 and 3 participate in degradation of extracellular matrix proteins.

The structure of the active site with three histidine residues for interaction with the catalytic zinc ion and Met-turn with the highly conservative sequence AL(V)MYP were identified in the protease domain of matrixins.

Matrilysin is the smallest protein among the matrixins, and other enzymes are mosaic proteins with additional domains at the C-end of the protease domain. Matrixins are synthesized as precursors with a signal peptide of 20 a.a. and propeptide containing up to 80 a.a. located at the N-end of enzyme molecule. The protease domain 160-170 a.a. in length contains ions of calcium and catalytically active zinc. Activation of the enzymes occurs via the "cysteine switch" mechanism on proteolytic cleavage of the pro-sequence by trypsin, plasmin, and other proteins [7, 18]. Comparison of amino acid sequences of protease domains of various matrixins demonstrated distinct homology indicating that these enzymes evolved from a common ancestor [30].

Disulfide bonds are absent from the protease domain of some collagenases (MMP-1, MMP-8), but several calcium ions necessary for stabilization of this domain and an additional noncatalytic zinc ion have been found; this zinc ion does not participate in stabilization of the structure and can be replaced by other metal ions [8, 18].

Collagenases MMP-1 and MMP-8 perform limited proteolysis of collagens of types I, II, and III, gelatins, proteoglycans, and other substrates [31]. The octapeptide Gly-Pro-Gln-Gly-Ile-Trp-Gly-Gln is optimal among synthetic substrates. The enzymes hydrolyze bonds in synthetic substrates if Ala is in P1 position and Tyr or Phe in P1′ position. Proteolytic activity of matrixins significantly increases in the presence of zinc (0.5 mM) and calcium (5 mM) ions and NaCl (100 mM) in the reaction mixture with neutral pH. α_2-Macroglobulin, some antibiotics, chelating agents, and also tissue inhibitors with low selectivity inhibit these enzymes [7].

Matrix metalloproteinases are able to decompose structural components of extracellular matrix and perform intracellular regulation of various pathological processes. They promote metastasis of tumor cells via destruction of matrix barriers, aggravation of chronic diseases, collagen decomposition in rheumatoid arthritis, and development of hypopyon ulcer, inflammation processes including enteral and atherosclerotic ones, meningitis, Alzheimer's disease, activation of various growth factors, and modeling of chemotaxis signals [29].

The family of adamalysins/reprolysins (MEROPS M12B) is formed by two groups of enzymes: adamalysins and reprolysins. Reprolysins were isolated from venom of various snakes and are called SVMPs (snake venom metalloproteinases). By now more than 200 kinds of snake venom are known; they contain 50-60 proteins and peptides with various structures and functions. Venom is composed of neurotoxins, cardiotoxins, various growth factors, disintegrins, bradykinin-potential proteins, enzymes (proteases, phospholipases, phosphodiesterases, cholinesterases, aminotransferases, catalase, ATPase, hyaluronidase, etc.), and protein inhibitors. Many reprolysins are isoenzymes isolated from a single snake species. So, from 13 reprolysins of *Crotalus atrax* venom half are isoenzymes whose amino acid sequences are 98% identical.

Reprolysins are multi-domain proteins containing a large set of non-peptidase domains at the C-end [32]. All reprolysins are divided into four classes by the number of domains and molecular mass [33, 34].

Class 1. Reprolysins with molecular mass 20-30 kDa, the mature form consists of one protease domain. Representatives: adamalysin II, atrolysins B and C, acutolysins A, B, and C.

Class 2. Reprolysins with molecular mass 30-50 kDa having a disintegrin domain for interaction with cell receptors along with protease domain (atrolysin E, bilitoxin, acutolysin D).

Class 3. Reprolysins and most hemorrhagic toxins with molecular mass 50-80 kDa composed of protease, disintegrin, and cysteine-containing domains (atrolysin A).

Class 4. Reprolysins with molecular mass 80-100 kDa containing a fourth lectin-like domain (metalloproteinase from *Vipera russelli* venom).

Some enzymes of class 1 are called "hemorrhagic toxins", and they are divided into two subclasses: subclass 1A with distinct hemorrhagic activity and subclass 1B with weak or null activity. Thus, adamalysin II from *Crotalus adamanteus* does not exhibit any hemorrhagic activity, atrolysin C has weak hemorrhagic activity, and both belong to subclass 1B, whereas acutolysin A has high hemorrhagic activity and is called a hemorrhagic proteinase. The pH optimum for proteolytic activity of these enzymes is in weakly basic medium (pH 7.5) and is 100 times higher than at pH 5.0. Chelating agents inhibit proteolytic and hemorrhagic activities [32]. Besides catalytically important zinc ion, these enzymes have a calcium ion on the surface of the molecule; Ca^{2+} ion binds to oxygen atoms thus providing structural stability of the catalytic domain [33]. Reprolysins are synthesized and stored in venom glands of snakes as inactive zymogens. The enzymes have three cysteine residues, and one of them participates in activation of these endopeptidases via the "cysteine switch" mechanism [34]. Substrate specificity of reprolysins is defined via hydrolysis of casein, B-chain of insulin, and fibrinogen. These enzymes perform limited proteolysis of natural substrate via one or two peptide bonds. Thus, trimerelysins 1 and 2 preferentially cleave α-chain of fibrinogen via the Pro516–Met517 bond, and a metalloproteinase from *Naja* asp venom performs limited proteolysis of α-chain of human fibrinogen via two bonds, Lys412–Leu413 and Phe501–Asp502, on extended incubation [32].

Reprolysins are mainly responsible for hemorrhagic effects and tissue necrosis from snake bites. Some reprolysins with high hemorrhagic activity (hemorrhagic toxins) are able to destroy cell membranes of capillaries, causing local blood effusion and destruction of components of extracellular matrix of victims and are responsible for cell adhesion [32].

A large number of endopeptidases isolated from reproductive tissues of mammals, mainly mice, rats, rabbits, and humans, belong to a group of adamalysins Their tertiary structure analogous to that of reprolysins looks like an ellipsoid form. They are characterized by the presence of a large set of additional domains located at the C-end of the protease domain (a unique disintegrin domain responsible for binding to integrin or other cell surface receptors, cysteine-rich domain, EGF-like domain, and others).

A signal sequence and pro-sequence are located at the N-end of adamalysins. The function of the pro-domain is to maintain the enzyme in the inactive state *in vivo*. Proteolytic excision of the propeptide activates the enzyme, and in some cases these enzymes can be subjected to autocatalytic activation. It was found that the pro-domain can act as a chaperon for correct folding of the protease domain and structuring of the active site [35].

The physicochemical properties of these enzymes are well studied. The activity optima of adamalysins are in a wide pH range 7.0-9.0. Fibronectin, laminin, gelatin, and collagen are substrates for adamalysins, chelating agents, non-selective peptide hydroxamates, and tissue inhibitors of metalloproteinases (TIMPs) are inhibitors. Adamalysins are activated via the "cysteine switch" mechanism, but their activation also may be related with conformational changes in the pro-domain [18, 35].

Adamalysins participate in conception processes, neurogenesis, proteolysis of cell surface proteins, and other physiological processes. ADAM-17 regulating activity of α-factor of tumor necrosis is the best-studied enzyme. Adamalysins are able to cleave and remodel the structure of extracellular matrix and matrix-associated growth factors, thus defining such physiological processes as migration and growth [36, 37]. Localization of adamalysins is different. It was found that one isoform of endopeptidase ADAM-12 is a membrane-bound protein and another is a secreted protein. It was also found that one isoform of human endopeptidase ADAM-28 is expressed in milt, whereas the second membrane-bound form is specific for lymphatic glands [35].

We isolated a new bacterial adamalysin-like metalloproteinase *B. intermedius* 3-19 (MprBi) from culture liquid of *B. subtilis* recombinant strain. The gene for the protein was cloned, and the nucleotide sequence was determined and registered in the International Gene Bank (number AN EU678894). Analysis of the *mprBi* gene sequence indicates that the signal peptide is located at the N-end of protein molecule; this is evidence for extracellular localization of the mature protein [38]. The enzyme was purified to homogeneity, and its primary structure was determined by MALDI-TOF spectrometry [39]. The N-terminal sequence of mature protein molecule – AST-GSQKVTV – was determined by the Edman method. As follows from this, alanine is the N-terminal amino acid in

Metzincin metalloproteinases	Active site motif												Met-turn				
ASTACINS																	
Astacin (crawfish)	**H**	**E**	L	M	**H**	A	I	**G**	F	Y	**H**	E	S	I	**M**	H	Y
α-MEP (mouse)	**H**	**E**	I	L	**H**	A	L	**G**	F	F	**H**	E	S	L	**M**	H	Y
β-MEP (rat)	**H**	**E**	F	L	**H**	A	L	**G**	F	W	**H**	E	S	V	**M**	H	Y
BMP1/procollagen C-proteinase (human)	**H**	**E**	L	G	**H**	V	V	**G**	F	W	**H**	E	S	I	**M**	H	Y
SPAN/BP10 (sea urchin)	**H**	**E**	I	G	**H**	A	I	**G**	F	H	**H**	E	S	I	**M**	H	Y
Tolloid-proteinase (*Dr. melanogaster*)	**H**	**E**	L	G	**H**	T	I	**G**	F	H	**H**	E	S	I	**M**	H	Y
Flavastacin (*F. meningosepticum*)	**H**	**E**	I	M	**H**	S	M	**G**	I	M	**H**	E	S	V	**M**	M	Y
SERRALYSINS																	
Serratia proteinase	**H**	**E**	I	G	**H**	A	L	**G**	L	S	**H**	P	S	L	**M**	S	Y
Proteinase B (*E. chrysanthemi*)	**H**	**E**	I	G	**H**	A	L	**G**	L	S	**H**	P	S	I	**M**	S	Y
P. aeruginosa proteinase [47]	**H**	**E**	I	G	**H**	T	L	**G**	L	S	**H**	P	S	V	**M**	S	V
MATRIXINS																	
MMP-1 (collagenase 1 of human fibroblasts)	**H**	**E**	L	G	**H**	S	L	**G**	L	S	**H**	S	A	L	**M**	Y	P
MMP-3 (human stromelysin-1)	**H**	**E**	I	G	**H**	S	L	**G**	L	F	**H**	S	A	L	**M**	Y	P
MMP-8 (neutrophilic collagenase 2)	**H**	**E**	F	G	**H**	S	L	**G**	L	A	**H**	S	A	L	**M**	Y	P
ADAMALYSINS/ REPROLYSINS																	
Adamalysin II (*C. adamanteus*)	**H**	**E**	L	G	**H**	N	L	**G**	M	E	**H**	D	C	I	**M**	R	P
Atrolysin C	**H**	**E**	L	G	**H**	N	L	**G**	M	E	**H**	D	C	I	**M**	R	P
Trimerelysin	**H**	**E**	L	G	**H**	N	L	**G**	M	E	**H**	D	C	I	**M**	S	D
Acutolysin A [39]	**H**	**E**	M	A	**H**	N	L	**G**	V	S	**H**	D	C	I	**M**	S	P
THERMOLYSIN																	
Thermolysin (*B. thermoproteolyticus*) [44]	**H**	**E**	L	T	**H**	A	V	T	D	Y	T	A	no				
	126	127	128	129	130	131	132	133	134	135	136	137	145	146	147	148	149
MprBi (*B. intermedius*)	**H**	**E**	Y	G	**H**	N	F	**G**	L	P	**H**	D	C	L	**M**	N	Y

Fig. 2. Marker conservative sequences of metzincin metalloendopeptidases: the extended motif of the active site and Met-turn.

the mature molecule. In the amino acid sequence of MprBi, we identified the 126**HE**YG**H**NF**G**LP**H**136 fragment that contains three conservative histidine residues His126, His130, and His136, glutamate residue E127 near the first histidine His126, and glycine residue Gly133 between the second and third histidine residues (Fig. 2). The structure of this fragment completely coincides with the conservative structure of the extended motif of the active site of metzincin endopeptidases (Fig. 2) [8]. In the structure of this protein we identified the only methionine residue Met147 and also Cys145 and Tyr149 residues in its nearest neighborhood; the latter are typical of various families of the clan of metzincins. The presence of these conservative residues indicates that the CLMNY fragment is the Met-turn sequence (Fig. 2). Two marker sequences in the primary structure of MprBi allow classification of the extracellular bacillar proteinase as an enzyme belonging to the clan of metzincins. The enzyme belongs to the adamalysins/reprolysins family, because in the active site motif of MprBi Asp is located after the third histidine (His136) and Cys145 is located in the Met-turn, which is typical only of enzymes of this family [18].

STRUCTURAL FEATURES OF METZINCINS

Tertiary structures of some representatives of the metzincin clan are well known. The astacin molecule is a commonly accepted model when structures of metzincins are studied. It has been shown that the form of the astacin molecule is a globule divided into two subdomains (N- and C-domains) by a deep, narrow, and long cleft, catalytically active zinc being on the bottom of this cleft. The structure of the N-domain consists of four bent wide parallel and one antiparallel β-sheets and two long α-helices [40].

The C-domain begins after the conservative glycine residue Gly99, it is less ordered than the N-domain, contains only one α-helix, and its size and form is different in various astacin-like enzymes.

In the structure of mature astacin there are conservative cysteine residues forming two disulfide bridges: one of them (Cys42–Cys198) binds the N- and C-domains, and the second (Cys64–Cys84) is located near the active site. These bridges define the net stability and rigidity of the protein molecule [41].

All enzymes belonging to the clan of metzincins and also thermolysins and thermolysin-like endoproteinases have similar zinc-binding regions; this suggests a common zinc-binding model in these proteins. In the active site motif of astacin, 92**H**ELM**H**AIGFY**H**102, the position of the first two zinc ligands, His92 and His96, is analogous to the position of the histidine residues in the active site of thermolysin **H**ELT**H**. This fact indicates that the catalytically active sites of astacin and thermolysin are similar.

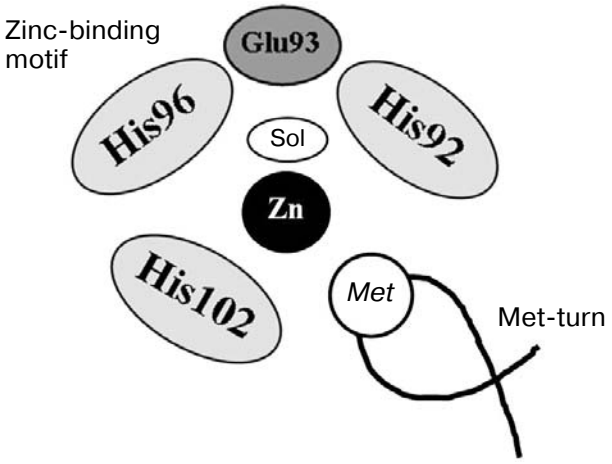

Fig. 3. Zinc-binding motif of the active site and Met-turn in the structure of the astacin molecule.

The conservative glycine residue Gly99 turns the polypeptide chain towards the zinc ion and plays an important role in active site folding. As a result, the imidazole ring of His102 is positioned opposite to the zinc ion and forms the third zinc ligand (Fig. 3) [40, 42].

Glutamic acid residue Glu93 located near the first histidine residue (His92) in the astacin molecule participates in catalysis as a catalytic base analogous to the catalytic process of thermolysin. To confirm this suggestion, a mutant protein with Glu93 replaced by Ala was obtained; this protein did not exhibit any proteolytic activity in hydrolysis of gelatin and highly sensitive fluorescent substrate. This experiment confirmed the status of Glu93 as a catalytic base in the astacin molecule, and H_2O becomes an additional fourth zinc ligand and participates in the catalytic act [43].

X-Ray structural analysis showed that in the mature astacin molecule the Glu103 residue positioned after the third histidine ligand is oriented with the involvement of a H_2O molecule so that a salt bridge with the N-terminal Ala1 residue is formed. Replacements of Glu103 by Gln or Ala negligibly influenced catalysis, but the enzyme became unstable. So, the Glu103–Ala1 bond is important for stabilization of the mature astacin molecule [40-42]. It should be mentioned that the Glu103 residue is typical of only the representatives of astacin family (Fig. 2). That is why Glu103 is sometimes also considered to be related with the **H**ELM**H**AIGFY**H***E* motif of the active site of astacin [40, 44, 45].

The Met-turn containing conservative methionine residue is present in the structures of all metzincins (Fig. 3). Methionine is a rare amino acid, and this is why its conservative presence in the Met-turn structure of metzincins is surprising [29]. As demonstrated by experiments, this amino acid does not play a key role in catalytic activity. Thus, the catalytic activity of endopeptidase

MMP2 (gelatinase A) negligibly changed on replacement of methionine in the Met-turn by leucine or serine [46]. In endopeptidase C from *Erwinia chrysanthemi* (PrtC) replacements of methionine by leucine, alanine, and isoleucine did not change stability, but the activity of the mutant proteins decreased. Crystallographic study of mutant proteins showed changes in geometry of the imidazole ring of the side chain, and it was suggested that these changes affect zinc binding to histidines of the active site [47]. Methionine in the Met-turn sequence of endopeptidase AprA from *P. aeruginosa* was replaced by L-difluoromethionine. Unlike leucine and isoleucine, L-difluoromethionine has minimal steric action due to two F atoms. This replacement caused a negligible change in catalytic activity of the mutant protein in comparison with the wild type protein, the thermal stability of the proteins being unchanged. The amino acid sequence of the Met-turn SI(L,V)MH**Y** of astacin-like enzymes is highly conservative, but only in astacins and serralysins was the tyrosine residue (Tyr149) found in this turn. Releasing a proton, Tyr149 forms a hydrogen bond with Zn and becomes the fifth zinc ligand [18]. It is supposed that a unique "tyrosine switch" can play a specific role in the catalytic process. Using site-directed mutagenesis, astacin with the Tyr149Phe replacement with low proteolytic activity was obtained. It was concluded that Tyr149 is not crucial for catalysis, but it can be responsible for enzyme binding to substrate and/or stabilization of the transition state during the catalytic act [43]. Study of absorption spectra of a Cu-astacin mutant provided support for participation of Tyr149 in substrate binding [48].

A two-stage mechanism of astacin activation was suggested in 2010. In the first stage of proenzyme activation, trypsin excises the pro-sequence of 34 amino acids ($Ala1^P$-$Asp34^P$), preparing the enzyme for transition to mature form. At this stage the enzyme becomes active. In the second activation stage, the enzyme structure is transformed; as a result the remaining six amino acids of propeptide ($Asp34^P$-$Ala1^M$) submerged in the protein molecule come to the surface of the protein globule and become accessible for cleavage. Subsequent release of the N-terminal Ala1 with formation of the mature molecule can occur either autocatalytically or be catalyzed by the same trypsin molecule [41].

For the families of matrixins and adamalysins/reprolysins, the fifth zinc ligand, a proton donor in catalysis, is unknown. In these proteins a Tyr residue in the Met-turn is mainly replaced by a non-functional proline residue, which does not participate in the catalytic act (Fig. 2). Propeptides of these metallopeptidases contain conservative fragments with a Cys residue that binds to Zn of the active site of the zymogen, blocking enzymatic activity. The enzyme is activated on proteolytic excision of the propeptide. A similar mechanism ("cysteine switch") was discovered in pro-domains of stromelysin-1 (pro MMP-3) and gelatinase A (pro MMP-2) and other enzymes of the matrixin family and also in enzymes of the adamalysins/reprolysins family. In the latter, cysteine in the composition of the conservative sequences PRCGVPD and PKMCGVT binds to zinc of the active site of the protease domain, maintaining the enzyme molecule in the inactive state. It is supposed that leishmanolysin with a pro-domain containing conservative sequence HRCIHD has the same mechanism of active site blocking [7].

So, metzincins are multi-domain proteins having a zinc-binding active site motif with three histidine ligands and a conservative Met-turn. The enzymes are activated via the mechanism of "tyrosine" or "cysteine switch". Participation of metzincins in various destructive and pathological processes influencing human health requires comprehensive study of their enzymatic properties and functional role for development of diagnostic tests and new drugs.

This work was financially supported by the federal program "Scientific and Pedagogical Personnel for Innovative Russia" for 2009-2013 (grant No. P406).

REFERENCES

1. Hoffmann, E. (2001) *Biochemistry (Moscow)*, **66**, 1144-1152.
2. Sterchi, E. E., Stocker, W., and Bond, J. S. (2008) *Molecular Aspects of Medicine*, **29**, 309-328.
3. *Enzyme Nomenclature 1992: Recombination of the Nomenclature Committee of International Union of Biochemistry and Molecular Biology and Classification of Enzymes* (1992) San Diego, p. 862.
4. Lopez-Otin, C., and Bond, J. S. (2008) *J. Biol. Chem.*, **283**, 30433-30437.
5. Rawlings, N. D., and Barrett, A. J. (1993) *Biochem. J.*, **290**, 205-218.
6. Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J., and Barrett, A. J. (2004) *Nucleic Acids Res.*, **32**, D160-D164.
7. Gomis-Ruth, F. X. (2003) *Mol. Biotechnol.*, **24**, 157-202.
8. Gomis-Ruth, F. X. (2009) *J. Biol. Chem.*, **284**, 15353-15357.
9. Pfleiderer, G., Zwilling, R., and Soneborn, H. H. (1967) *Hoppe-Seyler's Z. Physiol. Chem.*, **348**, 1319-1331.
10. Plummer, T. H., Jr., Tarentino, A. L., and Hauer, C. R. (1995) *J. Biol. Chem.*, **270**, 13192-13196.
11. Jiang, W., Gorbea, C. M., Flannery, A. V., Beynon, R. J., Grant, G. A., and Bond, J. S. (1992) *J. Biol. Chem.*, **267**, 9185-9193.
12. Gorbea, C. M., Marchand, P., Jiang, W., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Bond, J. S. (1993) *J. Biol. Chem.*, **268**, 21035-21043.
13. Yasumasu, S., Iuchi, I., and Yamagami, K. (1988) *Zool. Sci.*, **5**, 191-195.
14. Kessler, E., Adar, R., Goldberg, B., and Niece, R. (1986) *Collagen Rel. Res.*, **6**, 249-266.
15. Kessler, E., Takahara, K., Biniaminov, L., Brusel, M., and Greenspan, D. S. (1996) *Science*, **5271**, 360-362.

16. Yiallouros, I., Kappelhoff, R., Schilling, O., Wegmann, F., Helms, M. W., Auge, A., Brachtendorf, G., Berkhoff, E. G., Beermann, B., Hinz, H.-J., Konifg, S., Peter-Katalimic, J., and Stocker, W. (2002) *J. Mol. Biol.*, **342**, 237-246.
17. Titani, K., Torff, H.-J., Hormel, S., Kumar, S., Walsh, K. A., Rodl, J., Neurath, H., and Zwilling, R. (1987) *Biochemistry*, **26**, 385-392.
18. Stocker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Ruth, F.-X., McKay, D. B., and Bode, W. (1995) *Protein Sci.*, **4**, 823-840.
19. Bond, J. B., and Eynon, R. J. (1995) *Protein Sci.*, **4**, 1247-1261.
20. Tarentino, A. L., Quinones, G., Grimwood, B. G., Hauer, C. R., and Plummer, T. H., Jr. (1995) *Arch. Biochem. Biophys.*, **319**, 281-285.
21. Ge, G., and Greenspan, D. S. (2006) *Birth Defects Research (Pt. C)*, **78**, 47-68.
22. Morihara, K., Tsuzuki, H., and Oka, T. (1973) *Biochim. Biophys. Acta*, **309**, 414-429.
23. Marokhazi, J., Mihala, N., Hudecz, F., Fodor, A., Graf, L., and Venekei, I. (2007) *FEBS J.*, **274**, 1946-1956.
24. Letoffe, S., Delepelaire, P., and Wandersman, C. (1989) *Mol. Microbiol.*, **3**, 79-86.
25. Baumann, U. (1994) *J. Mol. Biol.*, **242**, 244-251.
26. Hamada, K., Hata, Y., Katsuya, Y., Hiramatsu, H., Fujiwara, T., and Katsube, Y. (1996) *J. Biochem.*, **119**, 844-851.
27. Miyatake, H., Hata, Y., Fujii, T., Hamada, K., Morihara, K., and Katsube, Y. (1995) *J. Biochem.*, **118**, 474-479.
28. Hege, T., and Baumann, U. (2001) *J. Mol. Biol.*, **314**, 187-193.
29. Walasek, P., and Honek, J. F. (2005) *BMC Biochemistry*, 6:21 doi:10.1186/147-2091-6-21.
30. Bode, W., Grams, F., Reinmer, P., Gomis-Ruth, F.-X., Baumann, U., McKay, D. B., and Stocker, W. (1996) *Adv. Exp. Med. Biol.*, **389**, 1-11.
31. Hasty, K. A., Jeffrey, J. J., Hibbs, M. S., and Welgus, H. G. (1987) *J. Biol. Chem.*, **262**, 313-318.
32. Matsui, T., Fujimura, Y., and Titani, K. (2000) *Biochim. Biophys. Acta*, **1477**, 146-156.
33. Gong, W., Zhu, X., Liu, S., Teng, M., and Niu, L. (1998) *J. Mol. Biol.*, **283**, 657-668.
34. Grams, F., Huber, R., Kress, L. F., Moroder, L., and Bode, W. (1993) *FEBS Lett.*, **335**, 76-80.
35. Seals, D. F., and Courtneidge, S. A. (2003) *Genes Dev.*, **17**, 7-30.
36. Blobel, C. P., and Apte, S. S. (2006) *Encyclopedia of Respiratory Medicine*, pp. 19-23.
37. Alfandari, D., Couin, H., Gaultier, A., Smith, K., White, J. M., Darribere, T., and DeSimone, D. W. (2001) *Curr. Biol.*, **11**, 918-930.
38. Sabirova, A. R., Rudakova, N. L., Balaban, N. P., Ilinskaya, O. N., Demiduyk, I. V., Kostrov, S. V., Rudenskaya, G. N., and Sharipova, M. R. (2010) *FEBS Lett.*, **584**, 4419-4425.
39. Rudakova, N. L., Balaban, N. P., Danilova, Yu. V., Rudenskaya, G. N., and Sharipova, M. R. (2010) *Biochemistry (Moscow)*, **75**, 1294-1301.
40. Bode, W., Gomis-Ruth, F.-X., Huber, R., Zwilling, R., and Stocker, W. (1992) *Nature*, **358**, 164-167.
41. Guevara, T., Yiallouros, I., Kappelhoff, R., Bissdorf, S., Stocker, W., and Gomis-Ruth, F.-X. (2010) *J. Biol. Chem.*, **285**, 13958-13965.
42. Bode, W., Gomis-Ruth, F.-X., and Stocker, W. (1993) *FEBS Lett.*, **331**, 134-140.
43. Yiallouros, I., Berkhoff, E. G., and Stocker, W. (2000) *FEBS Lett.*, **484**, 224-228.
44. Jiang, W., and Bond, J. S. (1992) *FEBS Lett.*, **312**, 110-114.
45. Gomis-Ruth, F.-X., Stocker, W., Huber, R., Zurilling, R., and Bode, W. (1993) *J. Mol. Biol.*, **229**, 945-968.
46. Butler, G. S., Tam, E. M., and Overall, C. M. (2004) *J. Biol. Chem.*, **279**, 15615-15620.
47. Oberholzer, A. E., Bumann, M., Hege, T., Russo, S., and Baumann, U. (2009) *J. Biol. Chem.*, **390**, 875-881.
48. Gomis-Ruth, F.-X., Grams, F., Yiallouros, I., Nar, H., Kusthardt, U., Zwilling, R., Bode, W., and Stocker, W. (1994) *J. Biol. Chem.*, **269**, 17111-17117.