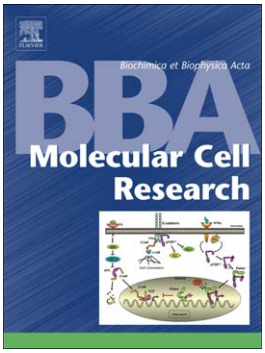

Matrix metalloproteinases outside vertebrates

Laura Marino-Puertas, Theodoros Goulas, F. Xavier Gomis-Rüth

PII: S0167-4889(17)30096-4
DOI: doi:10.1016/j.bbamcr.2017.04.003
Reference: BBAMCR 18072

To appear in: *BBA - Molecular Cell Research*

Received date: 13 February 2017
Revised date: 31 March 2017
Accepted date: 4 April 2017

Please cite this article as: Laura Marino-Puertas, Theodoros Goulas, F. Xavier Gomis-Rüth, Matrix metalloproteinases outside vertebrates, *BBA - Molecular Cell Research* (2017), doi:10.1016/j.bbamcr.2017.04.003

Matrix metalloproteinases outside vertebrates

Laura Marino-Puertas, Theodoros Goulas* and F. Xavier Gomis-Rüth*

Proteolysis Lab; Structural Biology Unit; "María-de-Maeztu" Unit of Excellence; Molecular Biology Institute of Barcelona (CSIC); Barcelona Science Park; c/Baldiri Reixac, 15-21; 08028 Barcelona (Spain).

* Corresponding authors: e-mail: thgcri@ibmb.csic.es, phone: (+34) 934 020 187 or e-mail: xgrcri@ibmb.csic.es, phone: (+34) 934 020 186.

ABSTRACT

The matrix metalloproteinase (MMP) family belongs to the metzincin clan of zinc-dependent metallopeptidases. Due to their enormous implications in physiology and disease, MMPs have mainly been studied in vertebrates. They are engaged in extracellular protein processing and degradation, and present extensive paralogy, with 23 forms in humans. One characteristic of MMPs is a ~165-residue catalytic domain (CD), which has been structurally studied for 14 MMPs from human, mouse, rat, pig and the oral-microbiome bacterium *Tannerella forsythia*. These studies revealed close overall coincidence and characteristic structural features, which distinguish MMPs from other metzincins and give rise to a sequence pattern for their identification. Here, we reviewed the literature available on MMPs outside vertebrates and performed database searches for potential MMP CDs in invertebrates, plants, fungi, viruses, protists, archaea and bacteria. These and previous results revealed that MMPs are widely present in several copies in Eumetazoa and higher plants (*Tracheophyta*), but have just token presence in eukaryotic algae. A few dozen sequences were found in *Ascomycota* (within fungi) and in double-stranded DNA viruses infecting invertebrates (within viruses). In contrast, a few hundred sequences were found in archaea and >1000 in bacteria, with several copies for some species. Most of the archaeal and bacterial phyla containing potential MMPs are present in human oral and gut microbiomes. Overall, MMP-like sequences are present across all kingdoms of life, but their asymmetric distribution contradicts the vertical descent model from a eubacterial or archaeal ancestor.

Keywords: zinc metalloproteinase; metzincin; MMP; invertebrates; catalytic domain; structure-based sequence motif

---

1. Molecular characteristics of matrix metalloproteinases **—** The matrix metalloproteinases (MMPs) are a widespread family of zinc-dependent metallopeptidases (MPs), which either broadly degrade extracellular matrix components or selectively activate or inactivate other proteins through limited proteolysis [1-3]. MMPs were discovered in 1962 as an active principle in frog metamorphosis [4], and they contain a central zinc-dependent catalytic domain (CD) of ~165 residues, which is mostly furnished at its N-terminus with a ~20-residue signal peptide for secretion and an ~80-residue zymogenic pro-domain (PD). Some MMPs possess a "furin recognition motif" (R-X-R/K-R; [5]) for proteolytic activation between the PD and the CD. Into this minimal configuration, distinct MMPs have inserted extra segments and domains, such as fibronectin-type-II inserts within the CD and C-terminal unstructured linker regions, ~200-residue hemopexin domains (HDs) and other domains, as well as glycosyl phosphatidylinositol (GPI) anchors and transmembrane segments [1, 3, 6]. Overall, this variability divides