Pergamon

Bioorganic & Medicinal Chemistry Letters 8 (1998) 853–858

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

STRUCTURAL INSIGHT INTO THE BINDING MOTIFS FOR THE CALCIUM ION AND THE NON-CATALYTIC ZINC IN MATRIX METALLOPROTEASES

Irina Massova, Lakshmi P. Kotra, and Shahriar Mobashery*

Department of Chemistry, Wayne State University, Detroit, MI 48202-3489, U.S.A.

Received 19 January 1998; accepted 23 February 1998

Abstract: The binding motifs for the structural zinc and calcium ions in matrix metalloproteases (MMPs) were investigated by analyzing the three-dimensional structural models of 23 representative MMPs.

Matrix metalloproteinases (MMPs) constitute a major group of proteases that carry out a myriad of physiological and pathological functions. These functions include embryogenesis, angiogenesis, wound healing, inflammation, arthritis, and cancer metastasis.[1,2] These enzymes are zinc-dependent endopeptidases known for their ability to cleave extra-cellular matrix constituents, as well as non-matrix proteins. MMPs have developed into a unique group of zinc-dependent proteinases characterized by the incorporation of various protein domains in their structures, which mediate interactions with substrates and inhibitors.[3] The implication of MMPs in cancer metastasis and angiogenesis has raised considerable interest in the MMP family since they represent attractive targets for development of novel anticancer drugs. Therefore, understanding of the structure and function of these key enzymes would have significant implications toward these efforts.

Figure 1. Stereo view of the ribbon drawing of the backbone of the catalytic domain of MMP-19. The red sphere represents the catalytic zinc ion, the green sphere represents the structural zinc ion and the orange sphere represents the calcium ion.

A total of 64 MMPs have been sequenced to date, of which 15 are from human. These human enzymes have counterparts in other vertebrates. Furthermore, MMPs have even been identified in invertebrates[4-6] and three have recently been sequenced from plant sources.[7] MMPs, in general, possess a propeptide domain (N-terminal), a catalytic domain and a hemopexin-like domain (C-terminal; except for MMP-7 which lacks it).[8] MMP-2 and -9 further possess a fibronectin-like domain and membrane-type MMPs (MT-MMPs) have acquired a transmembrane domain. In the catalytic domain, there are two zinc ions and at least one calcium ion coordinated to various residues (Figure 1). One of the two zinc ions is present in the active site, and is intimately involved in

0960-894X/98/$19.00
PII: S0960-894X(98)00128-0

the catalytic processes of these enzymes. The second zinc ion (structural zinc) and the calcium ion are present in the catalytic domain in proximity to the catalytic zinc (approximately 12 Å away). The catalytic zinc ion is essential for the proteolytic activity of MMPs and the three histidine residues that coordinate with the catalytic zinc are conserved among MMPs. However, there is little known about the roles of the second zinc ion and the calcium ion within the catalytic domain. It was noted that the MMPs have high affinities for the structural zinc and the calcium ions.[9,10] It has been suggested that these metal ions keep the structural elements together in the catalytic domains of the MMPs, contributing to its stability.[11-13] We have compared amino-acid sequences of 64 MMPs from various sources. Whereas the catalytic zinc and its coordinating residues are well conserved, structural zinc and the calcium ions show differences in their corresponding binding motifs. In an effort to understand the coordination patterns of the structural zinc and calcium ions in the catalytic domains of various MMPs, we have modeled the three-dimensional structures of several representative members of the MMPs. A clear insight into these coordination sites is important due to the fact that the structural zinc-binding motif as well as the calcium-binding motif are unique to the matrix metalloproteases and are not present in other metzincin proteases like astacin and adamalysin.[11] This comprehensive analysis of the three-dimensional models of the structural zinc and calcium ion binding regions has been performed here for the first time to expand our knowledge of this important family of enzymes.

Amino-acid sequences of MMPs were obtained from the GenBank, TREMBL, and SwissProt data banks. We utilized a total of 64 MMP sequences from various sources for the multiple-sequence alignment and then, a separate analysis was run using the representative 23 MMP sequences (*vide infra*). The human enzymes were selected where possible. Four human MMPs, fibroblast (MMP-1, 1cgl)[12] and neutrophil (MMP-8, 1mnc)[14] collagenases, matrilysin (MMP-7, 1mmq),[15] and stromelysin-1 (MMP-3, 1slm)[16] have recently been crystallized. The 17 modeled MMPs are the human MMP-10, MMP-11, MMP-12, MMP-13, MT1-MMP (MMP-14), MT2-MMP (MMP-15), MT3-MMP (MMP-16), MT4-MMP (MMP-17), MMP-19 (same as MMP-18), pig enamelysin, sea urchin envelysin (SwissProt accession number P22757), stromelysin-like MMP from newt (GenBank accession number D82053), collagenase-4 from frog, nematode MMP (GenBank accession number U00038), chicken CMMP, frog XMMP, and MMP from mouse-ear cress (TREMBL accession number O04529). The multiple-sequence alignments were performed using the program PileUp from the Wisconsin package version 9. We used the existing X-ray crystal structural information to predict the three-dimensional structures for the catalytic domains of 17 homologous metalloproteases using the program COMPOSER.[17] Water molecules that occupy the average positions in the crystallized MMPs and the two zinc and one calcium ions were added to the folded structures, and the entire complexes were allowed to undergo 10000 cycles of energy-minimization according to the methodology reported by us earlier.[18] We have reported recently on such computational structures for the catalytic domains of gelatinases A (MMP-2) and B (MMP-9).[18] After modeling the catalytic domains of the 17 MMPs, we utilized this information in conjunction with the four X-ray crystal structures of MMPs and previously modeled MMP-2 and MMP-9 structures to compare the structural zinc-binding and calcium-binding regions.

Figure 2 shows the multiple-sequence alignment for the structural zinc-binding and calcium-binding sites of the catalytic domains of the representative 23 MMPs. These metal-binding sites are formed by β-strands and turns. Analysis of the second zinc-binding site revealed that the MMPs developed at least four different motifs for

binding to this structural zinc ion. For the majority of MMPs (except for MMP-11, MMP-17, CMMP and XMMP), this site consists of the side chains of an aspartic acid and three histidine residues; which are consistent with the existing crystal structures. Two of the coordinating histidine imidazoles are bound as His_δ and one as His_ε (Figure 2; the three-dimensional structure of this site is shown in Figure 3A). The signature for this site in the 60 MMPs is the sequence(s) ***H***$_\delta$-[**G**N]-***D***-X(2)-[**P**AS]-**F**-***D***-[**G**A]-X(4)-[LIRV]-[**A**G]-***H***$_\delta$-[**A**V]-[**FYS**]-P-X(5,7,9)-***H***$_\varepsilon$-[**FL**]-***D***-X(2)-***E***-X-**W**, where residues that provide side chains for coordination to zinc and calcium are shown in bold italics, X (5,7,9) indicates 5, 7, or 9 variable residues in between the flanking sites (H_δ is a histidine protonated at $N_{\delta 1}$ and H_ε is a histidine protonated at $N_{\varepsilon 2}$). The first phenylalanine and the second phenylalanine/tyrosine (i.e., the FYS sequence) in this pattern are brought close to the coordinated histidines by the enzyme fold. This creates an increased hydrophobic environment, which presumably would enhance the binding affinity for the metal ions. The human MMP-11 has an aspartic acid instead of one of the His_δ. This is in contrast to the MMP-11 from rabbit and mouse, which still have a histidine at this position (data not shown). Therefore, we conclude that MMP-11 (compared to those from human, rabbit and mouse) does not have a unique motif for binding of the structural zinc.

```
MMP-1 human       DIMISFVRGDHRDNSP-FDGPGGNLAHAFQPGPGIGGD-AHFDEDERWTNN---------
MMP-2 human       DIMINFGRWEHGDGYP-FDGKDGLLAHAFAPGTGVGGD-SHFDDDELWTLGEGQVVRVKY
MMP-3 human       DIMISFAVREHGDFYP-FDGPGNVLAHAYAPGPGINGD-AHFDDDEQWTKD---------
MMP-7 human       DIMIGFARGAHGDSYP-FDGPGNTLAHAFAPGTGLGGD-AHFDEDERWTDG---------
MMP-8 human       DINIAFYQRDHGDNSP-FDGPNGILAHAFQPGQGIGGD-AHFDAEETWTNT---------
MMP-9 human       DIVIQFGVAEHGDGYP-FDGKDGLLAHAFPPGPGIQGD-AHFDDDELWSLGKGVVVPTRF
MMP-10 human      DIMISFAVKEHGDFYS-FDGPGHSLAHAYPPGPGLYGD-IHFDDDEKWTED---------
MMP-11 human      DIMIDFARYWDGDDLP-FDGPGGILAHAFFPKTHREGD-VHFDYDETWTIGDDQ------
MMP-12 human      DILVVFARGAHGDFHA-FDGKGGILAHAFGPGSGIGGD-AHFDEDEFWTTH---------
MMP-13 human      DIMISFGIKEHGDFYP-FDGPSGLLAHAFPPGPNYGGD-AHFDDDETWTSS---------
MT1-MMP human     DIMIFFAEGFHGDSTP-FDGEGGFLAHAYFPGPNIGGD-THFDSAEPWTVRNEDL-----
MT2-MMP human     DIMVLFASGFHGDSSP-FDGTGGFLAHAYFPGPGLGGD-THFDADEPWTFSSTDL-----
MT3-MMP human     DIPIIFASGFHGDSSP-FDGEGGFLAHAYFPGPGIGGD-THFDSDEPWTLGNPNH-----
MT4-MMP human     DIQIDFSKADHNDGYP-FDA-RRHRAHAFFPGHHHTAGYTHFNDDEAWTFRSSDA-----
MMP19 human       DIRLSFHGRQSSYCSNTFDGPGRVLAHADIPELG----SVHFDEDEFWTEGTY-------
Collagenase-4 frog DIEISFTAGDHKDNSP-FDGSGGILAHAFQPGNGIGGD-AHFDEDETWTKT---------
MMP newt          DIQISFGAREHGDFNP-FDGPYGTLAHAFAPGTGIGGD-AHFDEDEKWSKV---------
CMMP chicken      DIMVAFGTKAHGHCPRYFDGPLGVLAHAFPPGSGFGGD-VHFDEDEDWTMG---------
Enamelysin pig    DIMISFETGDHGDSYP-FDGPRGTLAHAFAPGEGLGGD-THFDNAEKWTMG---------
Envelysin         DIRIKFGSYDHGDGIS-FDGRGGVLAHAFLPRNG----DAHFDDSETWTEGTR-------
MMP nematode      DIYIAFEKGEHSDGFP-FDGQDGVVAHAFYPRDG----RLHFDAEEQWSLNSV-------
XMMP frog         DIKLGFGRGRHLGCSRAFDGSGQEFAHAWFLGD------IHFDDDEHFTAPS---------
MMP cress         DITIGFYTGDHGDGEP-FDGVLGTLAHAFSPPSG----KFHLDADENWVVSGDL------
```

Figure 2. Multiple-sequence alignment of the structural zinc and calcium binding regions in the catalytic domains of 23 representative MMPs. The structural zinc binding residues in majority of MMPs are given in blue - three histidines and one aspartic acid; the structural zinc binding domain in MMP-19, CMMP, and XMMP are in green. The residues that coordinate to the calcium ion are in pink; "underlined" amino acids chelate to the calcium ion via their backbone carbonyl moiety; the variation which was observed in the calcium binding site of MMP-17 (MT4-MMP) is given in cyan.

The remaining MMPs, those of chicken CMMP, the frog XMMP, and the human MMP-19 (Figure 2), each possess unique motifs for binding to the structural zinc ion. The chicken CMMP has a histidine where the

majority of MMPs have an aspartic acid coordinated to the second zinc ion. Consequently, the structural zinc ion in this enzyme is coordinated by four histidine residues (His-171, His-173, His-187 and His-200; Figure 3B). Such coordination would force the cysteine adjacent to the fourth coordinated histidine (Cys-174) to point to the outside milieu. This cysteine may provide a site for protein dimerization, as it is entirely exposed and available. Such active dimer formation has been reported for MMP-9, for example.[19] We also investigated the possibility for the existence of another binding mode for the second zinc ion, where the three original histidines and the cysteine are coordinated to the zinc ion (Figure 3C). Such structural alternative is a distinct possibility, and it would still be a novel motif for coordination to the structural zinc. The frog XMMP has a cysteine residue at the same position as in the sequence of CMMP. However, it is missing the aspartic acid at the usual location typical for most MMPs for coordination to the second zinc ion. This enzyme has a glycine in the place of aspartic acid, which obviously cannot provide metal coordination (Figure 2). Therefore, for XMMP the only possibility is to have three histidines (His-270, His-286 and His-294) and the cysteine (Cys-273) side chains to coordinate to the second zinc ion (Figure 3D).

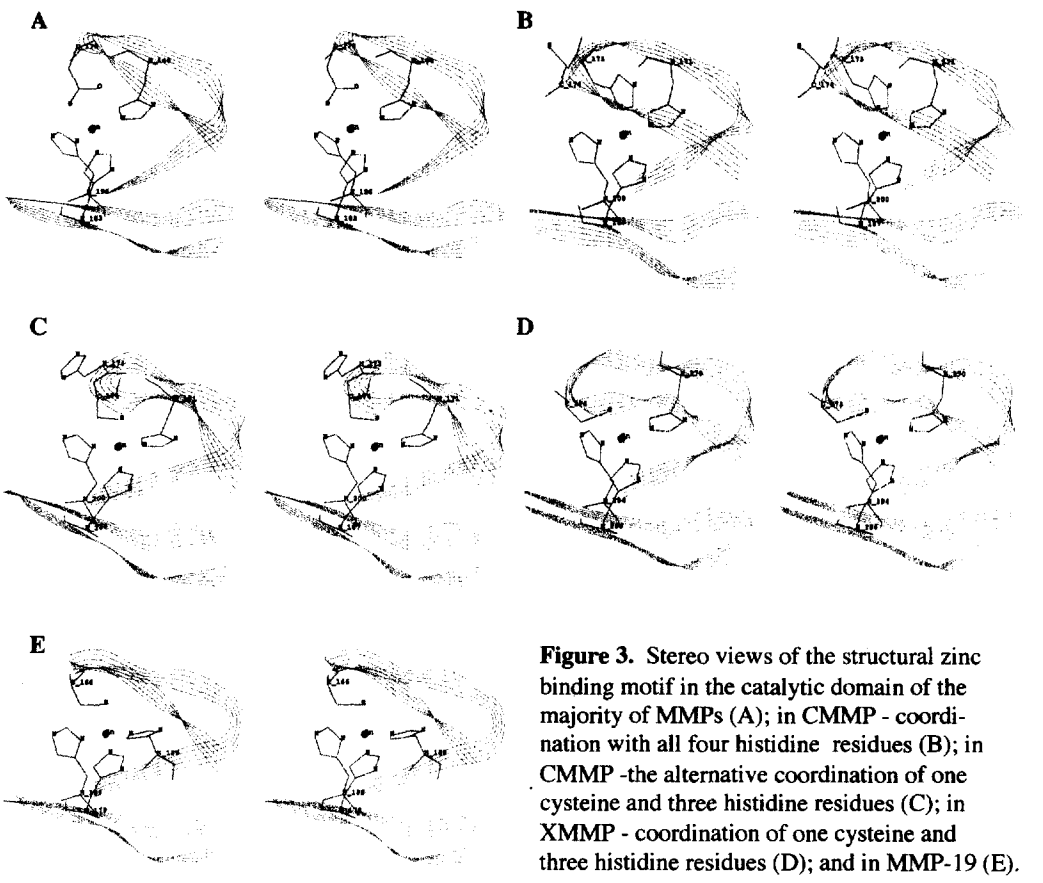

Figure 3. Stereo views of the structural zinc binding motif in the catalytic domain of the majority of MMPs (A); in CMMP - coordination with all four histidine residues (B); in CMMP -the alternative coordination of one cysteine and three histidine residues (C); in XMMP - coordination of one cysteine and three histidine residues (D); and in MMP-19 (E).

Analysis of the structure for MMP-19 showed that this enzyme has two histidines for metal coordination at the usual places and one cysteine at the corresponding position already discussed in the sequences for CMMP and XMMP. However, MMP-19 possesses a non-coordinating serine (Ser-163) at the position where the majority of MMPs have the third histidine. Interestingly, the third coordinating histidine (His-159, Figure 3E) in this enzyme is provided at an entirely distinct position, four residues to the N-terminus. Whereas the position of the amino acid is different and indeed on a different β-strand, it provided the side chain for the fourth coordination site in a perfectly acceptable orientation in space (for example, compare the positions of His-159 and His-270 in Figures 3E and 3D, respectively).

Binding of the calcium ion brings six specific elements of MMP catalytic domain for coordination in an octahedral fashion. Three of these elements are the side chains of two aspartic acids and one glutamic acid, which are conserved in all 23 MMPs except in MT4-MMP (*vide infra*, Figure 2). The remaining three calcium ligands are provided by the backbone carbonyl oxygens of three residues within a turn made up of five amino acids (Figure 2). A typical calcium-binding motif is shown in Figure 4A. The only exception to this general picture would appear to be MT4-MMP (MMP-17), which has undergone one amino acid deletion in this turn (i.e., it has a four-amino acid turn). In human MMP-17 (MT4-MMP), an asparagine residue is observed instead of the second aspartic acid (Figure 2). The coordination to the calcium ion for MT4-MMP (MMP-17) is shown in Figure 4B.

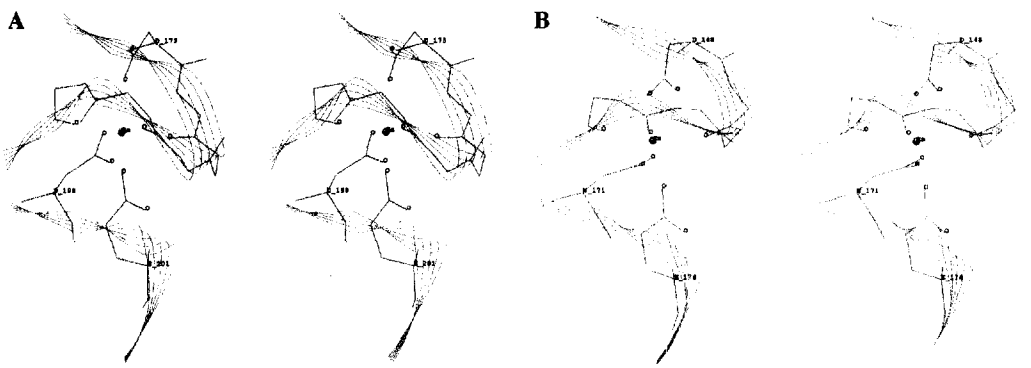

Figure 4. Stereo views of the calcium binding motif in majority of MMPs, except for MMP-17 (A); that for MMP-17 (B).

In conclusion, there are at least four different motifs for the binding to the structural zinc ion that are discernible from the three-dimensional structure analysis. However, the calcium binding motif is more strictly conserved. These variations in the metal-binding motifs preserve the topology of the structural zinc-binding site, as well as the calcium-binding site, and hence the general fold for the catalytic domain is highly preserved. The second zinc binding site is present in all MMPs and is important for the activity of these enzymes. Hence, the structural variations that are noted here in the coordination to the structural zinc ion may be indicative of the different outcomes for selection of novel enzymic activities. Furthermore, dendrogram analysis of the catalytic

domain sequence alignment for MMPs indicates that these variations in the zinc-binding motif came about as a consequence of independent evolutionary processes, unrelated to one another.[20] Since these motifs are absent in other members of the metzincin family of enzymes, we venture to say that the existence of these metal-binding sites in MMPs must have arisen in response to specific needs unique to MMPs with reference to their substrate specificities. The information provided herein is intended to stimulate interest in exploring the role of these motifs in MMPs geared toward understanding the substrate specificities for these enzymes, which are not understood to any appreciable degree at the present. Our understanding of the actual functions of the MMPs is at a rudimentary stage at the present and as more sequences of MMPs become available, the analysis presented here should be updated and correlated with the new structural information to shed light on the various properties of these important enzymes.

Acknowledgments. This work was supported by a grant from the US Army (DAMD17-97-1-7174). IM was a recipient of the Rumble and Heller predoctoral fellowships. We acknowledge Markku Kurkinen for providing us with the amino-acid sequence for CMMP.

References.

1. Chen, W. -T. *Curr. Opinion Cell Biol.* **1992**, *4*, 802.
2. Hagmann, W. K.; Lark, M. W.; Becker, J. W. In *Ann. Rev. Med. Chem.*; Bristol, J. A., Ed.; Academic Press, 1996; Vol. 31, pp 231-240.
3. Baramova, E.; Foidart, J. -M. *Cell Biol. Int.* **1995**, *19*, 239.
4. Lepage, T.; Gache, C. *EMBO J.* **1990**, *9*, 3003.
5. Namura, K.; Shimuzu, T.; Kinoh, H.; Sendai, Y.; Inomata, M.; Suzuki, N. *Biochemistry* **1997**, *36*, 7225.
6. Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, J.; Connell, M.; Copsey, T.; Cooper, J.; Coulson, A.; Craxton, M.; Dear, S.; Du, Z.; Durbin, R.; Favello, A.; Fraser, A.; Fulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hiller, L.; Jier, M.; Johnston, L.; Jones, M.; Kershaw, J.; Kirsten, J.; Laisster, N.; Latreille, P.; Lightning, J.; Lloyd, C.; Mortimore, B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D.; Shownkeen, R.; Sims, M.; Smaldon, N.; Smith, A.; Smith, M.; Sonnhammer, E.; Staden, R.; Sulston, J.; Thierry-Mieg, J.; Thomas, K.; Vaudin, M.; Vaughan, K.; Waterston, R.; Watson, A.; Weinstock, L.; Wilkinson-Sproat, J.; Wohldman, P. *Nature* **1994**, *368*, 32.
7. McGeehan, G.; Burkhart, W.; Anderegg, R.; Becherer, D.; Gillikin, J. W.; Graham, J. S. *Plant Physiol.* **1992**, *99*, 1179.
8. Murphy, G.; Knäuper, V. *Matrix Biol.* **1997**, *15*, 511.
9. Bode, W.; Reinemer, P.; Huber, R.; Kleine, T.; Schnierer, S.; Tschesche, H. *EMBO J.* **1994**, *13*, 1263.
10. Salowe, S. P.; Marcy, A. I.; Cuca, G. C.; Smith, C. K.; Kopka, I. E.; Hagman, W. K.; Hermes, J. D. *Biochemistry* **1993**, *31*, 4535.
11. Willenbrock, F.; Murphy, G.; Phillips, I. R.; Brocklehurst, K. *FEBS Lett.* **1995**, *358*, 189.
12. Lovejoy, B.; Cleasby, A.; Hassell, A. M.; Longley, K.; Luther, M. A.; Weigl, D.; McGeehan, G.; McElroy, A. B.; Drewry, D.; Lambert, M. H.; Jordan, S. R. *Science* **1994**, *263*, 375.
13. Wetmore, D. R.; Hardman, K. D. *Biochemistry* **1996**, *35*, 6549.
14. Browner, M. F.; Smith, W. W.; Castelhano, A. L. *Biochemistry* **1995**, *34*, 6602.
15. Gooley, P. R.; O'Connell, F.; Marcy, A. I.; Cuca, G. C.; Salowe, S. P.; Bush, B. L.; Hermes, J. D.; Esser, C. K.; Hagmann, W. K.; Springer, J. P.; Johnson, B. A. *Nature Struct. Biol.* **1994**, *1*, 111.
16. Stams, T.; Spurlino, J. C.; Smith, D. L.; Wahl, R. C.; Ho, T. F.; Qoronfleh, M. W.; Banks, T. M.; Rubin, B. *Nature Struct. Biol.* **1994**, *1*, 119.
17. Coordinates for the three-dimensional structures of the 17 MMPs that were modeled will be made available from our group web site, http://sun2.chem.wayne.edu/~somgroup upon acceptance of this manuscript.
18. Massova, I.; Fridman, R.; Mobashery, S. *J. Mol. Model.* **1997**, *3*, 17.
19. Clegg, P. D.; Burke, R. M.; Coughlan, A. R.; Riggs, C. M.; Carter, S. D. *Equine Vet. J.* **1997**, *29*, 343.
20. Massova, I.; Kotra, L. P.; Fridman, R.; Mobashery, S. *FASEB J.* submitted.