University of Florence

International Doctorate in Structural Biology

XXI Cycle (2006-08)

Matrix Metalloproteinase 12 (MMP12)/ Metalloelastase/Macrophage Elastase: Cloning, Expression, Structural & Interaction Studies

Ph.D. Thesis of

Rahul Jaiswal

Tutor

Dr. Marco Fragai

Coordinator

Prof. Claudio Luchinat

S.S.D. CHIM/03

This thesis has been approved by the University of Florence, the University of Frankfurt and the Utrecht University

I hereby declare that this thesis is based on the information extracted from the cited references and from the experimental work, which I have performed at Magnetic Resonance Center (CERM)

December 31, 2008 ……

Dedicated to my Parent, to my Nephew (Khushal), to All those who lost their Life in terror attack in Mumbai (November 26, 2008) and to *Escherischia coli*.....

RAHUL JAISWAL

Correspondence Address
Magnetic Resonance Center (CERM)
Via Luigi Sacconi, 6
University of Florence, Polo Scientifico
Sesto Fiorentino (Fi), 50019 ITALY
Email: jaiswal@cerm.unifi.it
Phone: +39 3298479430
+39 055 457 4245

Permanent Address
Ambika Bhawan
124, Shri Laxmi Nagar
Chotta Bangarda Road
Indore (M.P.), 452005 INDIA
Email: princenikhir@gmail.com
Phone: +91-9893906013
+91-731-2620425

PERSONAL INFORMATION

Date of Birth: June 24th, 1981
Sex: Male
Nationality: Indian

ACADEMIC RECORD:-

- Master of Science in Industrial Microbiology (July 2002-June 2004) with a GGPA 8.2/10 from Devi Ahliya University, Indore (INDIA) with six month dissertation (Jan.2004-June 2004) with Prof. K.V.R. Chary, Molecular Biophysics Group, Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai (INDIA)
  Thesis Title: Over expression of a Calcium Binding Protein (CaBP II) of *Entamoeba histolytica* and its structural studies (using CD Spectroscopy, MALDI Spectroscopy, DLS Spectroscopy, & NMR Spectroscopy),
- Bachelor of Science in Microbiology (July 1999-June 2002) from Devi Ahliya University, Indore (INDIA) with 65.5% aggregate.

RESEARCH EXPERIENCE:

- Doctoral Study: (Jan.2006-till date) with Dr. Marco Fragai and Prof. Claudio Luchinat at Magnetic Resonance Center (CERM), University of Florence, Florence (ITALY)
  Thesis Title: "Structural Characterization, activation mechanism and interaction studies of Matrix Metalloproteinase 12 (MMP 12)",
- Pre Doctoral Training: (April 2005-Dec. 2005) with Dr. Marco Fragai and Prof. Claudio Luchinat at Magnetic Resonance Center (CERM), University of Florence, Florence (ITALY)
  Title: "Expression, Purification and Structural Characterization of human matrix metalloproteinase (MMPs)",
- Project Trainee: (July 2004-Jan.2005): with Prof. K.V.R. Chary and Prof. Grijesh Govil, Molecular Biophysics Group, Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai (INDIA)
  Title: "Effect of L-Arginine on the metabolism and motility in the asthenospermic human spermatozoa".

FELLOWSHP/HONOR/AWARD:

- Carl Storm International Diversity Fellowship (Year 2007) in Gordon Research Conference on Matrix Metalloproteinase in Il Ciocco, Lucca, ITALY,
- Jan 2006-present: *Assegno di Ricerca* fellowship awarded by the University of Florence, ITALY,
- April 2005-Dec.2005: Fellowship awarded by CIRMMP, University of Florence, ITALY,
- July 2004-Jan. 2005: Project Associate Fellowship awarded by Department of Chemical Sciences, Tata Institute of Fundamental Research, Mumbai, INDIA,
- Qualified Graduate Aptitude Test in Engineering GATE conducted by Indian Institute of Technology (IIT) with an All India Ranking 150 in Life Sciences

PUBLICATIONS:

- Bertini, I., Calderone, V., Fragai, M., **Jaiswal, R.,** Luchinat, C., Melikian, M., Mylonas, E., Svergun, D. I. (2008) Evidence of reciprocal reorientation of the catalytic and hemopexin-like domains of full length MMP-12. *Journal of American Chemical Society, **130*** (22), 7011-21,
- Calderone, V., Dragoni, E., Fragai, M, **Jaiswal, R.,** Luchinat, C., Nativi, C. Biotin-Tagged Probes for Profiling of MMPs Expression and Activation: Design, Synthesis and Binding Properties (Submitted).

PS: Authors are ordered alphabetically by their surname.

CONFRENCE/POSTER:

- Gordon Research Conferences on Matrix Metalloproteinases, June 3-8, 2007, Il Ciocco, Barga (ITALY) (Poster Presentation),
- Advances and Management of NMR in Life Sciences, January 18-20, 2007, Florence, (ITALY) (Participant)
- International Conference of Magnetic Resonance in Biological Science (ICMRBS), January 16-20, 2005, Hyderabad, (INDIA) (Poster Presentation),
- International Congress of Chemistry and Environment, December 16-18, 2001, Indore, (INDIA), (Participant)

Acknowledgements

The three and half years of my life that I have spent during the course of my Graduation degree have been extremely insightful and pedantic. It is a great experience to be here and work here in Magnetic Resonance Center (CERM). Being in this institution has taught me both the "how" and "what" aspects of learning.

I am deeply indebted to my advisor Dr. Marco Fragai. He has been instructional and scholastic in his guidance as a Thesis Instructor. He has been a cool, caring and an encouraging guide not only during the entire course of my research work in the institute but also during my this long stay here in Florence city. I should submit that without his support it would not have been possible for me to complete my work.

I would like to thank Prof. Claudio Luchinat for his guidance, valuable suggestions and encouragement which helped me throughout my work. I would like to render my gratitude to Prof. Ivano Bertini, who has given me an opportunity to work in the CERM, University of Florence.

I would like to render my special gratitude to Dr. Anusarka Bhaumik and Antonella Nessi for their valuable suggestions and support throughout my work and stay from the beginning to the completion and also in compiling my work in the form of a thesis and for being patient listeners during excruciating times.

This work which bears my name would not be possible single handedly, if it were not for the people who constantly helped me immensely in every aspect of my existence in the institution. I would like to thank Dr. Leonardo Gonnelli, Dr. Marco Allegrozzi, Dr. Kwan Joo Yeo, Dr. Manuele Migliardi, Dr. Chiara Venturi, Mirco Toccafondi, Maxime Melikian, Niko Sarti and each and every student of CERM for all their timely help and cooperation. I would like to thank deep from my heart to all the faculties, technicians and Administrative & Secretariat office persons of CERM.

I extend my heartfelt thanks to my roomies and colleagues Dr. Anusarka Bhaumik, Soumyasri Das-Gupta, and Dr. Ravi Krishnan Elangovan for their wonderful company,

support and friendship. My special thanks to Dr. Maheswaran Shanmugam, Dr. Murugendra Vanarotti, D.S.S. Hembram, Shailesh Sharma, Malini Nagulapalli and Vaishali Sharma who makes my stay in Florence easy and enjoyable. I also express my thanks through this to my house owner Mr. and Mrs. Sandretti for their affection and care.

I express a deep sense of gratitude to my dearest friends Fatema Husain, Ajit Muley, Alka Bikrol, Anil Verma, Omesh Nandanwar, Sandeep Bhatia, Gaurav Pathak, Milind Maheshwari and Mubeen A. Khan for their love, affection and encouragement through course of my work though thousands miles away.

Last but not the least; I am indebted to my father, mother, sisters (Rashmi and Deepa), Brother-in-law (Jeeju, Sanjay) and my genius little champ, my nephew (Khush) for standing by my side ALWAYS... in good times and especially in bad ones. Their love and encouragement has contributed immensely to this work beyond words can express.

Contents

ABBREVIATIONS

1. MMPs: Matrix Metalloproteinase.
2. ECM: Extracellular Matrix.
3. Hpx: Hemopexin-like.
4. Cat: Catalytic.
5. NMR: Nuclear Magnetic Resonance.
6. CD: Circular Dichroism.
7. SPR: Surface Plasmon Resonance.
8. PDB: Protein Data Bank.
9. AHA: AcetoHydroxamate.
10. NOE: Nuclear Overhauser effect.
11. RDCs: Residual Dipolar Couplings.
12. ITC: Isothermal Titration Calorimetry.
13. BTI: Biotin Tagged Inhibitor.
14. HSQC: Heteronuclear Single Quantum Coherence.
15. IEC: Ion-exchange Chromatography.

1. Aims and Topics of the Research

Aims and topics of research

Matrix metalloproteinases (MMPs) are zinc dependent endopeptidases capable of degrading all kinds of extracellular matrix (ECM) proteins under normal physiological conditions. As MMPs have long been associated with many pathological conditions and in many stages of cancer, this protease family is a potential pharmaceutical target. Many inhibitors have been developed to therapeutically block the ECM degrading activities of MMPs in metastasis and angiogenesis.

The recent approach to drug design is to understand the molecular determinants of specificity and selectivity. The structural information of the catalytic mechanism of MMPs is of great importance, as this information may provide a better insight of the entities involved in the interaction/inhibition and useful in designing new effective drug targets for this class of proteases.

Aiming to address the question, the present research has been carried out to understand the molecular mechanism of catalysis of MMPs using MMP12 (or macrophage elastase) as a model, because of the high level of structural and functional homology within members of family. This project was focused on investigating the mechanism, at the molecular level, not only within the active site but also to its downstream or upstream sites and to design new candidate drugs with increased selectivity and specificity based on such structural and molecular information. In this context Full-length MMP12 and its catalytic and hemopexin-like domains were cloned and expressed in a high throughput fashion and structurally characterized utilizing various biochemical and biophysical techniques.

One of the major problems to work with MMPs is their instability because of their high activity and auto-proteolysis property. A major part of the work was to standardize strategies and protocols to obtain pure protein with higher stability at the given conditions for longer duration required for NMR or other biophysical characterization. In this context, different mutants for catalytic domain and full-length have been cloned and successfully expressed in *E.coli*. Further, during the course of the work the protocol has been standardized to replace the catalytic zinc metal ion with an equivalent metal ion like

cadmium(II) or the paramagnetic metal ion cobalt(II) without affecting the structural conformation but with higher stability.

Structurally, MMPs are multi-domains proteins. The flexibility of various domains in the protein and their ability to rearrange in the presence of the substrate represent a serious challenge against specificity. To accommodate the substrate in the active site, there is a necessity of relative mobility of the domains in the protein. In the attempt to understand the catalytic mechanism, a part of project has been devoted to address the hypothesized reorientation of the catalytic and hemopexin-like domains. The three dimensional structure of full-length MMP12 was solved by X-ray; the structure of the hemopexin-like domain was solved by NMR, while the structure of the catalytic domain was available both from X-ray and NMR. In particular the behavior of the catalytic and hemopexin-like domains in solution has been investigated to reveal the reciprocal mobility of these domains of MMP12. This finding might be relevant for other members of family having a shorter linker region between the two domains.

The other fascinating area of research to understand the catalysis process is to investigate the substrate-enzyme recognition mechanism. In MMPs the catalytic domain alone bears full proteolytic activity towards a range of substrates. It is often hypothesized that the hemopexin-like domain helps in substrate/ligand recognition and binding and results in efficient proteolysis. However, for MMP12 it has been reported that the catalytic domain alone is able to degrade elastin without the hemopexin-like domain. To understand how such group of noncollagenase MMPs recognize and bind the substrate and elicit proteolysis is of particular importance in determining their physiological roles. In this context, the analysis of interaction of MMP12 with elastin can provide relevant information and might be fruitful to address the long debated role of hemopexin-like domain for this protease. Aiming to address this question, the project has recently developed in the direction of studying the interaction of individual domains as well as of full-length MMP12 with elastin and its components. Further, an attempt has been made to clone and express tropoelastin, the building block of elastin, in order to verify the molecular determinants of substrate recognition.

In silico prediction of the structure of protein-ligand adducts through docking programs can provide valuable information in the early ligand design phases but cannot be used to

analyze the subtleties of interactions. By interpreting detailed structural information with a thermodynamic analysis of the ligand binding it is possible to clarify the driving force responsible for the affinity. In this way it is possible to plan modifications of the ligand able to increase its affinity and selectivity. In this respect we have solved the crystal structure of MMP12 in complex with a selective inhibitor and characterized the interaction in solution by NMR and microcalorimetry.

Another fascinating aspect of drug design is the design of molecular probes able to selectively bind biomolecular targets. This approach is gaining importance not only for its many possible applications in molecular biology, but also for the development of new techniques for the early stage diagnosis and therapy of several pathologies. MMPs expression and activation is increased in almost all human cancers comparing to physiological conditions. It is therefore apparent that the development of efficient molecular devices designed to reveal the overexpression of MMP may be a powerful tool for the early detection of tumors. In this aspect, utilizing the extensive structural and biochemical information that we have gained on MMP12, a part of the present work involved the design and synthesis of carboxylic-based biotinylated inhibitors able to show high affinity for several MMPs and at the same time able to strongly interact with avidin through the biotinylated tag.

The present thesis is based on the following articles:

- Bertini, I., Calderone, V., Fragai, M., **Jaiswal, R.,** Luchinat, C., Melikian, M., Mylonas, E., Svergun, D. I. (2008) Evidence of reciprocal reorientation of the catalytic and hemopexin-like domains of full length MMP-12. *J Am Chem Soc,* ***130*** (22), 7011-21,

- Calderone, V., Dragoni, E., Fragai, M, **Jaiswal, R.,** Luchinat, C., Nativi, C. Biotin-Tagged Probes for Profiling of MMPs Expression and Activation: Design, Synthesis and Binding Properties (Submitted to *Bioconjug. Chem.*),

- Structural and Thermodynamic analysis of a selective inhibitor of Matrix Metalloelastase (In *preparation*).

2. INTRODUCTION

2.1 Extracellular Matrix (ECM)

In biology, the extracellular matrix (ECM) is the extracellular part of animal tissues that usually provides structural support to the cells in addition to performing various other important functions. The extracellular matrix is the defining feature of connective tissue in animals. Extracellular matrix includes the interstitial matrix and the basement membrane [(1)]. Interstitial matrix is present between various cells (i.e., in the intercellular spaces). Gels of polysaccharides and fibrous proteins fill the interstitial space and act as a compression buffer against the stress placed on the ECM [(2)] (Fig. 1). Basement membranes are sheet-like depositions of ECM on which various epithelial cells rest.

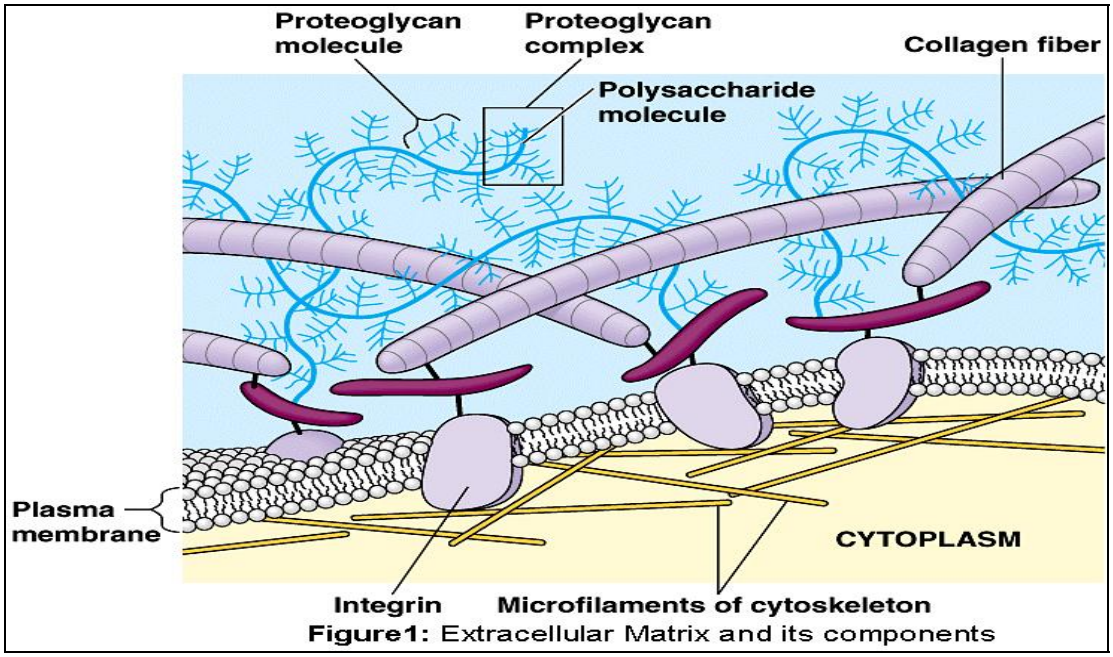


Figure1: Extracellular Matrix and its components

Role and Importance:

Due to its diverse nature and composition, the ECM can serve many functions, such as providing support and anchorage for cells, segregating tissues from one another, and regulating intercellular communication. The ECM regulates a cell's dynamic behavior. In addition, it sequesters a wide range of cellular growth factors, and acts as a local depot for them. Changes in physiological conditions can trigger protease activities that cause local release of such depots. This allows the rapid and local growth factor-mediated activation of cellular functions, without de novo synthesis.

An understanding of ECM structure and composition also helps in comprehending the complex dynamics of tumor invasion and metastasis in cancer biology as metastasis often involves the destruction of extracellular matrix [3].

Molecular components

Components of the ECM are produced intracellularly by resident cells, and secreted into the ECM via exocytosis [4]. Once secreted then they aggregate with the existing matrix. The ECM is composed of:

Proteoglycan matrix components

Glycosaminoglycans are carbohydrate polymers and are usually attached to extracellular matrix proteins to form proteoglycans. Proteoglycans have a net negative charge that attracts water molecules, keeping the ECM and resident cells hydrated.

Heparan sulfate proteoglycans, Chondroitin sulfate proteoglycans, Keratan sulfate proteoglycans are various types of proteoglycans.

Non-proteoglycan matrix components

Hyaluronic acid, Collagen, Fibronectin, Elastin and Laminin

Cell adhesion to the ECM

Many cells bind to components of the extracellular matrix. This cell-to-ECM adhesion is regulated by specific cell surface cellular adhesion molecules (CAM) known as integrins. Integrins are cell surface proteins that bind cells to ECM structures, such as fibronectin and laminin, and also to integrin proteins on the surface of other cells.

Cell types involved in ECM formation

There are many cell types that contribute to the development of the various types of extracellular matrix found in plethora of tissue types. The local components of ECM determine the properties of the connective tissue.

Fibroblasts are the most common cell type in connective tissue ECM, in which they synthesize, maintain and provide a structural framework; fibroblasts secrete the precursor components of the ECM, including the ground substance. Chondrocytes are

found in cartilage and produce the cartilagenous matrix. Osteoblasts are responsible for bone formation.

2.2 Matrix Metalloproteinases (MMPs)

Matrix Metalloproteinases, abbreviated as MMPs are family of zinc dependent endopeptidases that in normal physiological conditions are responsible for the regulation of extracellular matrix. MMPs were initially discovered in 1962 as Gross and Lapiere reported that tadpole tails during metamorphosis contained an enzyme that could degrade fibrillar collagen [(5)]. They are included in the "MB clan" of metallopeptidases, containing the motif HEXXHXXGXXH as the zinc-binding active site. The MB clan members are generically referred to as "Metzincins" since they all contain a conserved methionine that forms a turn eight residues downstream from the active site. Clan MB contains a number of families, and MMPs are in family M10. Family M10 is further divided into subfamilies A and B, and MMPs are in subfamily A, also known as the "matrixins." MMPs from vertebrate species are given MMP numbers (i.e. MMP1, MMP2, etc.), and MMPs from invertebrates rely on trivial names. The MMP family consists of at least 26 members, all of which share a common catalytic core with a zinc molecule in the active site.

The MMPs are produced as zymogens, with a signal sequence and propeptide segment that must be removed during activation. The propeptide domain contains a conserved cysteine, which chelates the active zinc site. An exception is MMP-23 which lacks this conserved cysteine, and has a very different propeptide domain. The generally conserved sequence PRCGVP around the chelating cysteine has been called the "cysteine switch." A subset of MMPs, including the membrane bound MMPs (MT-MMPs), as well as MMP- 11, MMP-21, MMP-23, and MMP-28, contain a basic prohormone convertase cleavage sequence (RRKR, RRRR, RKRR, etc.), which is thought to be cleaved by the PACE/Furin family enzymes. Two of the MMPs (MMP-2 and MMP-9) contain a fibronectin-like domain inserted into the catalytic domain, presumably to enhance substrate binding. MMP-9 also contains a collagen type-V-like domain, which may enhance substrate binding and specificity. MMP-21 from Xenopus also contains a vitronectin-like region just after the cysteine switch, presumably to assist in substrate interaction.

Though, initially classified as proteinases capable of digesting structural components of ECM, MMPs proteolytic targets expanded to several other extracellular matrix proteins. This wide range of substrates includes many other proteases, protease inhibitors, clotting factors, chemotactic molecules, latent growth factors, growth factor binding proteins, cell surface receptors, and cell-cell and cell matrix adhesion molecules [6]. Functional regulation of MMPs occurs at multiple levels. MMP mRNA expression is under tight, cell type– dependent control, with expression of individual MMPs associated with specific inflammatory, connective tissue, or epithelial cell types. MMP transcripts are generally expressed at low levels, but these levels rise rapidly when tissues undergo remodeling, such as in inflammation, wound healing, and cancer. MMPs can be categorized into various subgroup depending on the structure or function basis.

2.3 Structural and Functional Classification of MMPs

Earlier, MMPs were divided into subgroup on the basis of their specificity for the ECM components and thus reflects in the common name e.g., Collagenases, Gelatinases, Stromelysins and matrilysins (Fig.2).

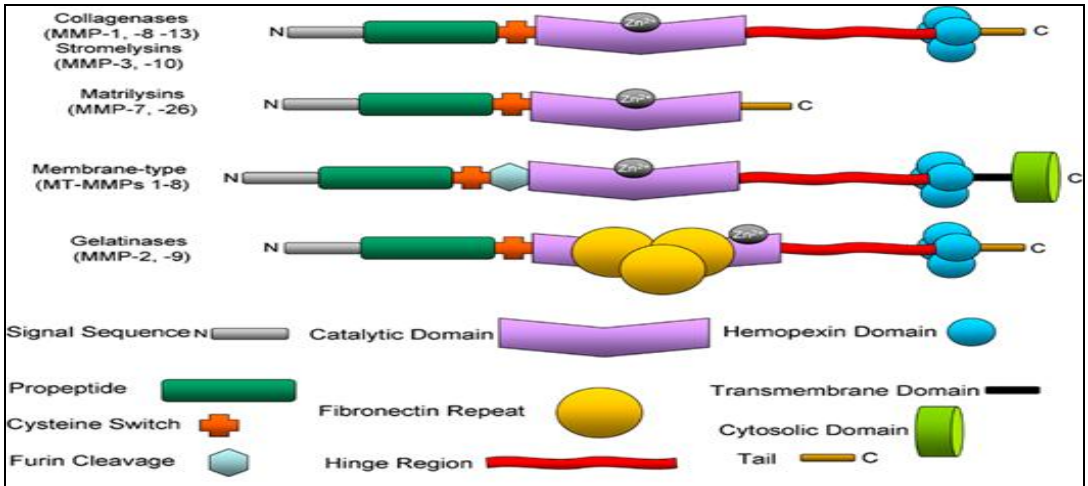


Fig.2 MMP structure classified according to the substrates they degrade and possess several general common structural characteristics. Adapted from A K Chow, J Cena and R Schulz *British Journal of Pharmacology* (2007) **152,** 189–205

All MMPs have an N-terminal signal sequence (or "pre" domain) that is removed after it directs their synthesis to the endoplasmic reticulum. Thus most MMPs are secreted; however, six display transmembrane domains and are expressed as cell surface enzymes. The pre domain is followed by a propeptide "pro" domain that maintains

enzyme latency until it is removed or disrupted, and a catalytic domain that contains the conserved zinc-binding region [7]. The catalytic domain dictates cleavage site specificity through its active site cleft, through specificity sub-site pockets that bind amino acid residues immediately adjacent to the scissile peptide bond, and through secondary substrate-binding exosites located outside the active site itself [8]. With the exception of MMP7 (matrilysin), MMP26 (endometase/matrilysin-2), and MMP23, all MMPs have a hemopexin/vitronectin-like domain that is connected to the catalytic domain by a hinge or linker region. MMP7 and MMP26 merely lack these extra domains, whereas MMP23 has unique cysteine-rich, proline-rich, and IL-1 type II receptor-like domains instead of a hemopexin domain [9, 10].

With the increasing range of substrates, MMPs are now sequentially numbered and grouped according to their structures. On the structural basis MMPs are categorized in 8 classes, five of which are secreted and three are membrane-type (MT) MMPs (Fig.3).

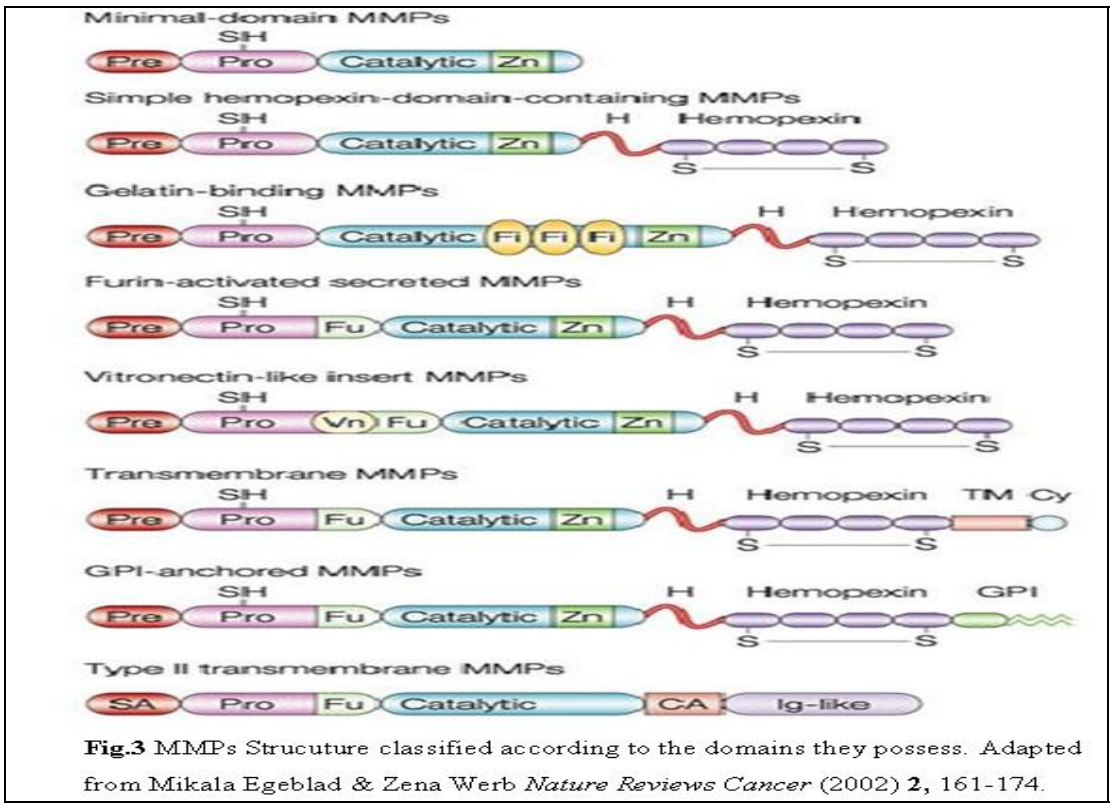


Fig.3 MMPs Strucuture classified according to the domains they possess. Adapted from Mikala Egeblad & Zena Werb *Nature Reviews Cancer* (2002) **2,** 161-174.

2.4 MMPs: Normal physiological regulation and Diseases

The pericellular MMPs are believed to be the physiological mediators of matrix degradation or turnover and, collectively, the MMPs can cleave all protein

components of the extracellular matrix (ECM) [11]. Indeed, many peptides liberated by the partial proteolysis of ECM macromolecules can regulate cell activities. Besides the ECM, other substrates of MMP processing include growth factors, receptors and adhesion molecules [12].

These activities confer on MMPs functions as diverse as cellular differentiation, migration, regulation of growth factor activity, survival or apoptosis, angiogenesis, inflammation and signaling [13, 14]. The normal possible modes of action of MMPs are depicted in Fig.4. The simplicity of thinking about MMPs solely as extracellular matrix degrading enzymes has been eroded by the recognition of numerous other important roles for these proteins.

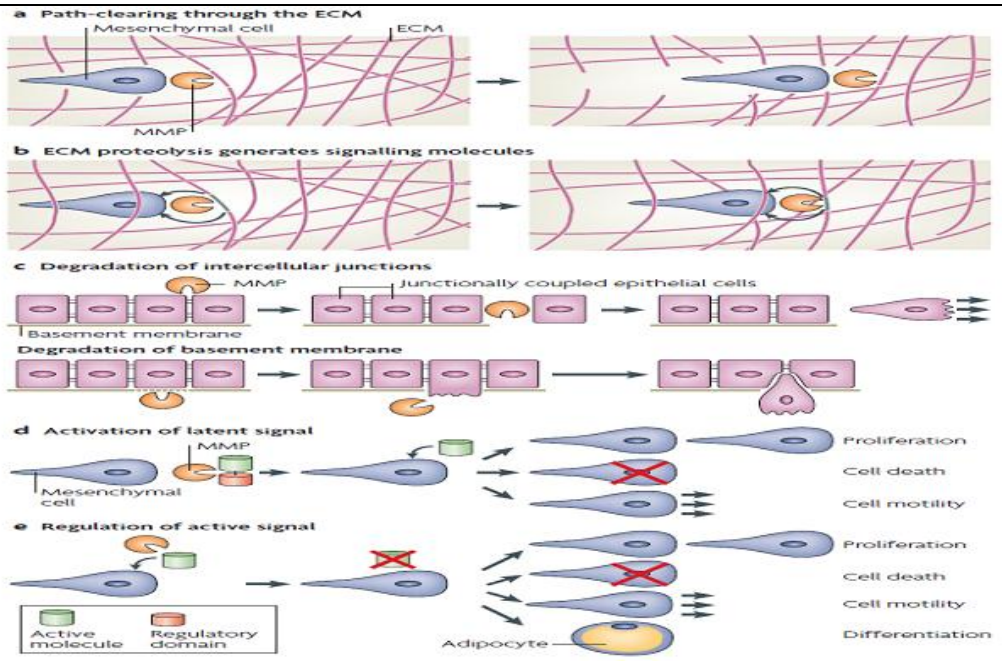


Fig.4 Possible modes of MMP action: a) MMPs can cleave component of ECM, resulting in increased space for cell or tissue movement. b) Alternatively, MMPs proteolysis can generate specific cleavage products that then signal in autocrine or paracrine manner. c) MMPs can also directly regulate the epithelial tissue architecture through cleavage of intracellular junctions or the basement membrane. d) MMPs can activate or modify the action of latent signaling molecules, resulting in diverse cellular consequences. e) MMPs can deactivate or modify the action of active signaling molecules, resulting in changes in proliferation, cell death, differentiation or cell motility. Adapted from Andrea Page-McCaw, Andrew J. Ewald, Zena Werb *Nature Review Molecular Cell Biology* (2007) **8**, 221-233.

Participation of the MMPs in various aspects of Cancer

MMPs are collectively able to degrade virtually all ECM components. In cancer, special emphasis was initially placed on the degradation of type IV collagen, a major protein component of basement membranes by MMP2 and MMP9. Subsequently, it

has been demonstrated that many non-ECM proteins can be cleaved by selected MMPs. Deciding which protein substrates are physiologically and pathologically relevant has been difficult to ascertain.

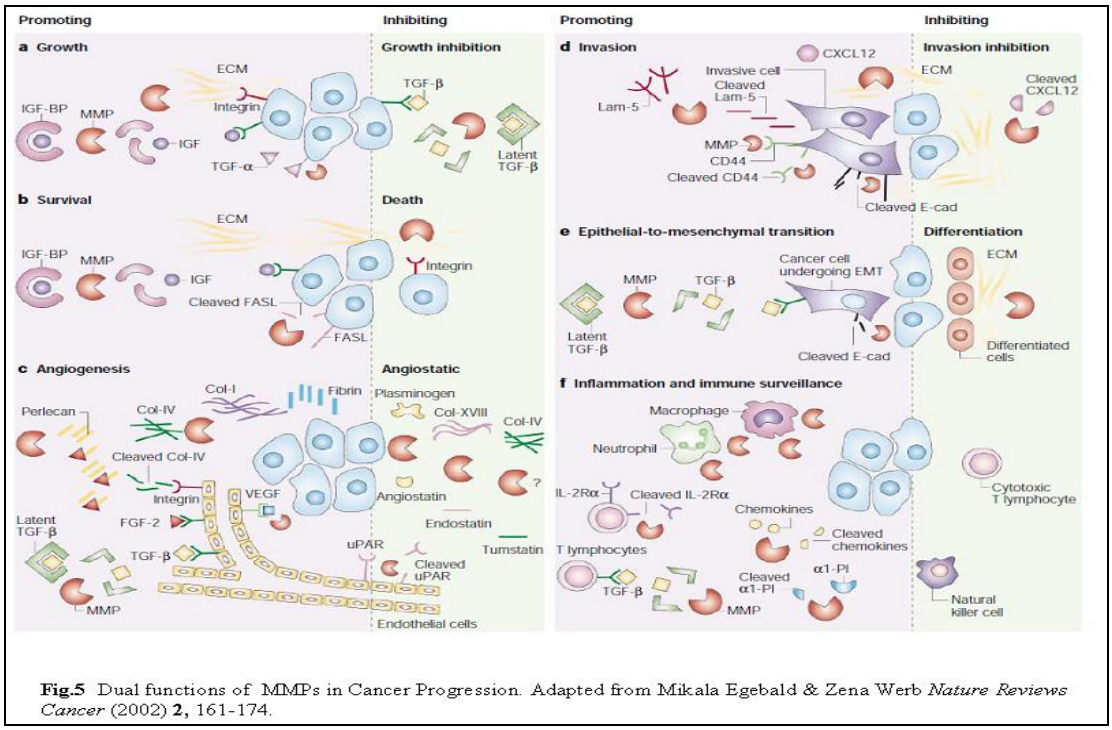


Fig.5 Dual functions of MMPs in Cancer Progression. Adapted from Mikala Egebald & Zena Werb *Nature Reviews Cancer* (2002) **2,** 161-174.

MMP-induced release from the cell surface (shedding) of heparin binding epithelial growth factor, insulin-like growth factor, and fibroblast growth factor enhance cell proliferation. On the other hand, release and activation of ECM sequestered TGFß by MMPs can lead to inhibition of cell proliferation. MMPs are capable of inducing both positive and negative effects on angiogenesis and immunity (Fig. 5). It has been observed that some MMPs readily cleave specific chemokines which may have profound effects on numerous biologic processes, not just limited to inflammation [15, 16].

Most MMPs in tumors are produced by stromal cells rather than the cancer cells (Fig. 6). One explanation for this phenomenon is that cancer cells produce Extracellular Matrix Metalloproteinase Inducer (EMMPRIN), a cell surface glycoprotein, which directly stimulates fibroblasts (through direct cell contact) to produce set of MMPs [17]. The importance of cytokines such as TNF-α, interleukin (IL)-1, and IL-6 in stimulating production of MMPs in disease has been emphasized.

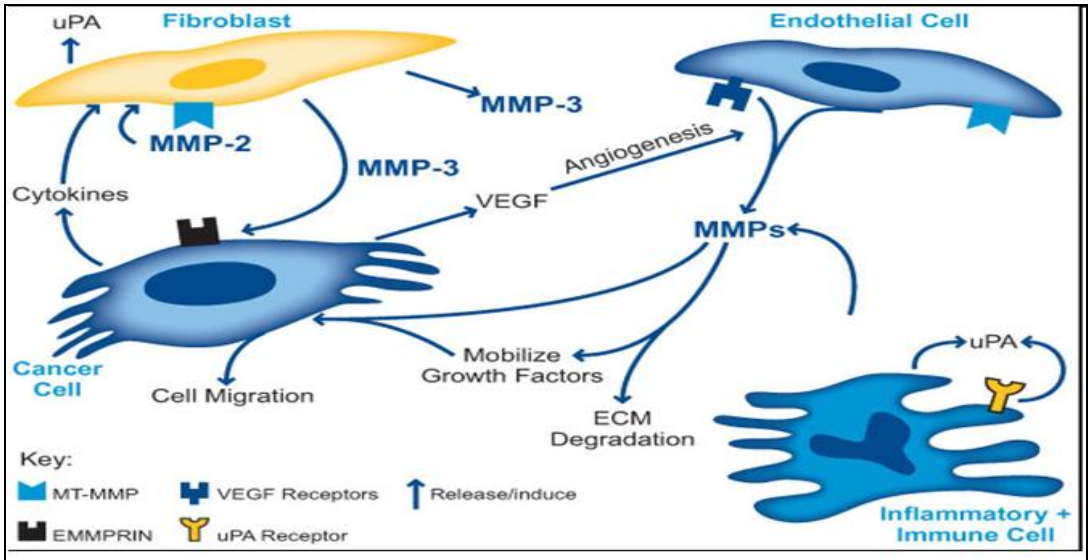


Fig.6 Schematic diagram of interaction between cancer cells and stromal within a tumor with a focus on role of MMPs. Adapted from Jian Cao & Stanley Zucker Biology and Chemistry of MMPs

Inflammatory diseases

Many reports have implicated MMP1, MMP3, and MMP9 in rheumatoid and osteoarthritis. An important role for aggrecanase, a member of the ADAM family of metalloproteinases, in articular damage has been proposed. A clinical trial of an MMPI in arthritis failed to achieve its goal.

Cardiovascular disease

There has been a long-standing interest in the role of MMPs in cardiovascular disease. Numerous studies have demonstrated increased levels of MMPs at sites of atherosclerosis and aneurysm formation [(18)]. The concept that the inflammatory process may play a leading role in the development of atherosclerotic plaques, has led to the suggestion that secretion and activation of MMPs by macrophages induces degradation of extracellular matrix in the atherosclerotic plaque and plaque rupture.

Lung disease

Elevated levels of MMPs have been implicated in the pathophysiology of various lung diseases, including acute respiratory distress syndrome, asthma, bronchiectasis, and cystic fibrosis. MMPs, EMMPRIN, and TIMPs are produced by many of the resident cells in the lung, hence complicating the analysis of their role in disease.

Central Nervous System disease

Following observations of the critical role of MMP9 in animal models resembling multiple sclerosis and Guillain-Barre's syndrome, MMPs has been implicated in several different types of neurologic diseases [19]. Treatment with synthetic inhibitors of MMPs has reversed some of the pathology in animal models of brain injury, especially stroke.

Shock syndromes

MMP8 and MMP9 are stored in the granules of polymorphonuclear leukocytes. These cells are key effectors in inflammatory and infectious processes. A role for these MMPs in shock is supported by studies in MMP9 deficient mice that were shown to be resistant to endotoxic shock [20].

Chronic wounds and inflammation of the skin and oral cavity

Acute and chronic wounds are associated with high levels of MMP2 and MMP9. These observations have led to the suggestion that nonhealing ulcers develop an environment containing high levels of activated MMPs, which results in chronic tissue turnover and failure of wound closure. MMP9 has been implicated in blistering skin diseases and contact hypersensitivity [21]. MMPs have long been implicated in periodontal disease [22] and more recently, in inflammatory bowel diseases.

2.5 Matrix Metalloproteinase 12 or Metalloelastase

MMP12 also known as Macrophage Elastase or Metalloelastase is a member of matrix metalloproteinase (MMP) gene family was first detected by Werb and Gordon in 1975 when they identified the elastolytic activity in mouse peritoneal macrophage conditioned media [23]. Like other members of MMPs family, MMP12 shares many features typical of MMPs including the domain structure, chromosomal location in MMP gene cluster in human chromosome 11q22, and to degrade components of ECM [24]. As other MMPs, MMP-12 is composed of three distinct domains: an amino-terminal propeptide domain that is involved in the maintenance of enzyme latency; a catalytic domain that binds zinc and calcium ion and hemopexin-like domain at the carboxy terminal which determines substrate specificity. The human gene, which is designated human macrophage metalloelastase, produces a 1.8-kb transcript encoding a 470-amino acid protein (Fig. 7)

that is 64% identical to the mouse protein. MMP-12 is secreted as a 54-kd pro-form protein that undergoes self-activation through autolytic processing [25]. Both the mRNA and protein were detected in alveolar macrophages.

Fig.7 Amino acid sequence of full length human metalloelastase (PDB Id: NP_002417). The leader sequence (underlined), the Cysteine Switch (bold face, orange colored), and the zinc binding active site (boxed, bold face, green colored).

MMP-12 is unique with respect to its predominantly macrophage-specific pattern of expression and the ability to readily shed its carboxyl-terminal domain upon processing. It degrades a broad range of ECM proteins, including elastin, type IV collagen, fibronectin, laminin and gelatin [26], and is involved in turnover of the matrix, cell migration, tissue repairing and remodeling. In addition, MMP-12 can activate other MMPs, for example, MMP-2 and -3, leading to subsequent degradation of other ECM proteins [27]. MMP-12, like other MMPs, also cleaves a variety of non-ECM proteins, such as plasminogen [28] and latent tumor necrosis factor α (TNFα), resulting in angiostatin and active TNFα, [29] respectively. Because of its great ability to degrade ECM, it is physiologically synthesized and released by macrophages to penetrate basement membranes and to invade normal and diseased tissue. The extensive role of Metalloelastase both in physiological and pathological conditions makes the enzyme an important target for medical research.

2.6 MMP12 in diseases

MMP12 was first detected in alveolar macrophages of cigarette smokers, but since then MMP-12 production has also been shown in other cell types and organs, e.g. in brain tumors [30] and placenta [31]. Elastin serves an important function in arteries and is particularly abundant in large elastic blood vessels such as the aorta. Elastin is also very important in the lungs, elastic ligaments, the skin, the bladder, elastic cartilage, and the intervertebral disc above the sacroiliac. Abnormal expression of MMP12 in these tissues

thus results in various pathological conditions in particular to those associated with the respiratory tract [32].

Emphysema

Chronic obstructive pulmonary disease (COPD) is a major global healthcare problem with tremendous morbidity and high mortality rates. Emphysema is defined as enlargement of peripheral airspaces of the lung accompanied by destruction of the walls of these structures [33]. The pathological hallmarks of the disease are small airway inflammation and obstruction and destruction of the ECM of the lung parenchyma that eventually lead to emphysema [34]. Chronic cigarette smoke (CS) exposure, which is associated with chronic inflammation and oxidative stress within lungs [35], is regarded as the primary cause of emphysema [36]. It is characterized by an accumulation of inflammatory cells such as macrophages and neutrophils. Indeed, it has been shown that cigarette smoke consistently produces an increase in the neutrophil number in bronchoalveaolar lavage fluid and in tissue [37, 38, 39]. Macrophage numbers are also elevated in the lungs of smokers and patients with COPD where they accumulate in the alveoli, bronchioli and small airways. Furthermore, there is a positive correlation between macrophage number in the alveolar walls and the mild-to-moderate emphysema status in patients with COPD [40]. Studies in mice demonstrated that macrophage-mediated lung alveoli destruction (emphysema) following cigarette smoke exposure is directly linked to the presence of MMP-12 [41] (Fig. 8). Thus, macrophage elastase is required for both macrophage accumulation and emphysema resulting from chronic inhalation of cigarette smoke.

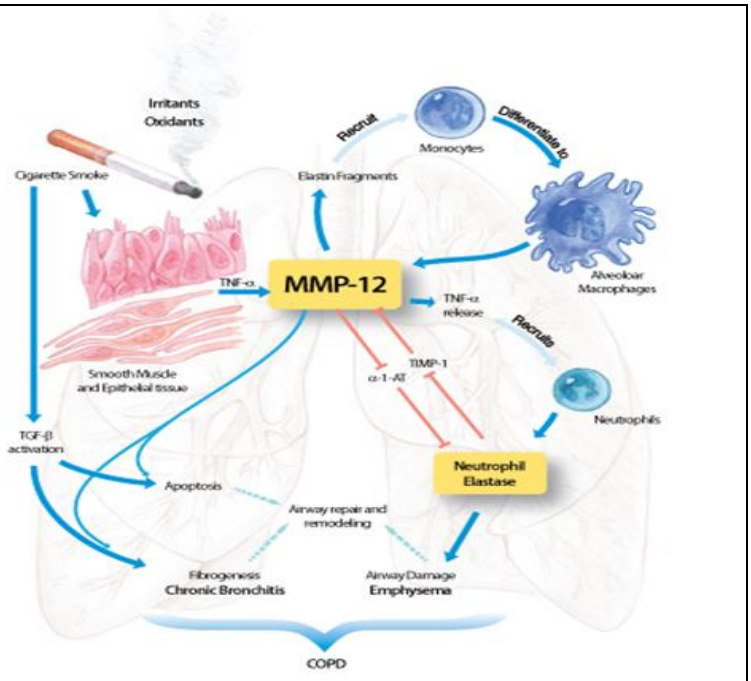


Fig.8 MMP12 and its proinflammatory properties. Adapted from http://www.rndsystems.com/cb_detail_objectname_sp07_MMP12.aspx

Vascular Aneurysms

Abdominal aortic aneurysms (AAA) affect between 2% and 9% of adults older than 65 years of age. AAAs is associated with aging, atherosclerosis, cigarette smoking, and male gender. AAAs is believed to result from structural remodeling of the elastin-rich aortic wall in association with chronic transmural inflammation. Mononuclear cell infiltration in the aortic media associated with elastic fiber destruction might be of particular importance. MMP-9 [(42)] and MMP-12 [(43)] were found to be prominently expressed in macrophages associated with elastic fiber disruption in specimens of human abdominal aortic aneurysm. Production of macrophage elastase was increased seven-fold in patients with aneurysms, and unlike other proteinases, expression was localized to the active "transition zone" in the media. This is the region adjacent to normal aorta where there is active remodeling and elastin degradation. Unlike other MMPs, macrophage elastase also specifically bound to fragmented, but not intact, elastic fibers in this zone. In vitro studies confirmed that macrophage elastase has a greater binding affinity for elastin than other elastases. These findings suggest that macrophage elastase might play a central role in aneurysm formation in humans.

Cancer

MMPs are believed to promote tumor progression by initiating carcinogenesis, enhancing tumor angiogenesis, disrupting local tissue architecture to allow tumor growth, and breaking down basement membrane barriers for metastatic spread. Several serine [44, 45] and metalloproteinases are capable of generating angiostatin [46] and other antiangiogenic plasminogen cleavage products, such as kringle domain 5 [47]. Generation of angiostatin in primary LLC tumors correlated with the presence of macrophages and macrophage elastase [48, 49]. With respect to MMPs, MMP-12 has the greatest proteolytic capacity to generate angiostatin from plasminogen. In fact, in the presence of plasminogen, wild-type murine macrophage conditioned medium acquires the capacity to inhibit endothelial cell proliferation by more than 50%. This effect is not observed in MMP-12-deficient macrophages. The importance of MMP-12 in limiting lung metastasis growth in the LLC model has been confirmed by use of mice deficient in macrophage elastase (MMP-12-/-) by gene-targeting. Studies suggest that local expression of MMP-12 in macrophages surrounding lung metastases limit growth. This effect does not appear to be related to generation of angiostatin. MMP12 is also expressed in malignant cells such as skin cancer, [50] astrocytomas, glioblastomas, [51] and hepatocellular cancer, in which its expression is associated with hypovascularization [52, 53]. This might be caused by the ability of MMP12 to cleave plasminogen to angiostatin and other kringle products, leading to the inhibition of angiogenesis in these tumors. [54, 55].

2.7 References

1. Kumar, V., Abbas, A.K., Fausto, N. (2005) *Robbins and Cotran: Pathologic Basis of Disease*; Elsevier; 7th edition.
2. Alberts B., Bray D., Hopin, K., Johnson, A., Lewis J, Raff, M., Roberts, K., Walter, P. (2004) Tissues and Cancer *Essential cell biology,* New York and London: Garland Science.
3. Liotta, L.A., Tryggvason, K., Garbisa, S., Hart, I., Foltz, C.M., Shafie, S. (1980) Metastatic potential correlates with enzymatic degradation of basement membrane collagen. *Nature* **284** (5751): 67–68.
4. Plopper, G. (2007) The extracellular matrix and cell adhesion, *Cells*. Sudbury, MA: Jones and Bartlett.
5. Gross, J., Lapiere, C. M. (1962) Collagenolytic activity in amphibian tissues: a tissue culture assay. *Proc. Nat.l Acad. Sci. USA* **47**, 1014–1022.
6. McCawley, L.J., Matrisian, L.M (2001) Matrix metalloproteinases: they're not just for matrix anymore! *Curr. Opin. Cell Biol.* **13**, 534.
7. Nagase, H., Woessner, J.F. Jr. (1999) Matrix metalloproteinases. *J. Biol. Chem.* **274**, 21491–94.
8. Overall, C.M. (2001) Matrix metalloproteinase substrate binding domains, modules and exosites. *Matrix Metalloproteinase Protocols*, ed. IM Clark, 79–120. Totowa, NJ: Humana.
9. Gururajan, R., Grenet, J., Lahti, J.M., Kidd, V.J. (1998) Isolation and characterization of two novel metalloproteinase genes linked to the CdC2Llocus on human chromosome 1p36.3. *Genomics* **52**:101–6.
10. Park, H.I., Ni, J., Gerkema, F.E., Liu, D., Belozerov, V.E., Sang, Q.X. (2000) Identification and characterization of human endometase (matrix metalloproteinase-26) from endometrial tumor. *J. Biol. Chem.* **275**, 20540–44.
11. Mott, J. D., Werb, Z. (2004) Regulation of matrix biology by matrix metalloproteinases. *Curr. Opin. Cell Biol.* **16**, 558–564.
12. McCawley, L. J., Matrisian, L. M. (2001) Matrix metalloproteinases: they're not just for matrix anymore! *Curr. Opin. Cell Biol.* **13**, 534–540.
13. Sternlicht, M. D., Werb, Z. (2001) How matrix metalloproteinases regulate cell behavior. *Annu. Rev. Cell Dev. Biol.* **17**, 463–516.

14. Parks, W. C., Wilson, C. L. Lopez-Boado, Y. S. (2004) Matrix metalloproteinases as modulators of inflammation and innate immunity. *Nature Rev. Immunol.* **4**, 617–629.
15. Zucker, S., Pei, D., Cao, J., Lopez-Otin, C. (2003) Membrane type-matrix metalloproteinases (MT-MMP). in: *Cell Surface Proteases*. Eds. S. Zucker and W-T Chen. Academic Press. 1-74.
16. Overall, C. M., Lopez-Otin, C. (2002) Strategies for MMP inhibition in cancer: innovations for the post-trial era. *Nature Reviews/Cancer.* **2**, 657-672.
17. Yan, L., Zucker, S., Toole, B. (2005) Roles of the multifunctional glycoprotein, emmprin (basigin; CD147), in tumour progression. *Thromb. Haemost.* **93** (2), 199–204.
18. Van den Steen, P.E., Dubois, B., Nelissen, I., Rudd, P.M., Dwek, R.A., Opdenakker, G. (2002) Biochemistry and molecular biology of gelatinase B or matrix metalloproteinase-9 (MMP-9). *Crit Rev Biochem Mol Biol.* **37**, 376-536.
19. Kieseier, B.C., Clements, J.M., Pischel, H.B., Wells, G.M., Miller, K., Gearing, A.J., Hartung, H.P. (1998) Matrix metalloproteinases MMP-9 and MMP-7 are expressed in experimental autoimmune neuritis and the Guillain-Barré syndrome. *Ann Neurol.* **43** (4) 427-34.
20. Dubois, B., Starckx, S., Pagenstecher, A., Oord, J., Arnold, B., Opdenakker, G. (2002) Gelatinase B deficiency protects against endotoxin shock. *Eur J Immunol.* **32** (8) 2163-71.
21. Liu, Z., Li, N., Diaz, L.A., Shipley, M., Senior, R.M., Werb, Z. (2005) Synergy between a plasminogen cascade and MMP-9 in autoimmune disease. *J Clin Invest.* **115** (4) 879-87.
22. Birkedal-Hansen, H. (1993) Role of matrix metalloproteinases in human periodontal diseases. *J Periodontol.* **64** (5 Suppl):474-84.
23. Werb, Z., Gordon, S. (1975) Elastase secretion by stimulated macrophages. *J Exp Med*. **142**, 361-377.
24. Shapiro, S.D. (1998) Matrix metalloproteinase degradation of extracellular matrix: biological consequences. *Curr Opin Cell Biol* **10**, 602-608.
25. Shapiro, S.D., Kobayashi, D.K., Ley, T.J. (1993) Cloning and characterization of a unique elastolytic metalloproteinase produced by human alveolar macrophages. *J Biol Chem* **268**, 23824-23829.

26. Gronski, T.J. Jr., Martin, R.L., Kobayashi, D.K., *et al*. (1997) Hydrolysis of a broad spectrum of extracellular matrix proteins by human macrophage elastase. *J Biol Chem* **272**, 12189–121894.
27. Matsumoto, S., Kobayashi, T., Katoh, M., *et al.* (1998) Expression and localization of matrix metalloproteinase-12 in the aorta of cholesterol-fed rabbits: relationship to lesion development. *Am. J. Pathol.* **153**, 109–119.
28. Cornelius, L.A., Nehring, L., Klein, B., Pierce, R., Bolinski, M., Welgus, H.G., *et al*. (1998) Generation of angiostatin by matrix metalloproteinases: effects on neovascularlation. *J Immunol*. **161** 6845-6852.
29. Chandler, S., Cossins, J., Lury, J., Wells, G. (1996) Macrophage metalloelastase degrades matrix and myelin proteins and processes a tumor necrosis factor-alpha fusion protein. *Biochem Biophys Res Commun*. **228**, 421-429.
30. Cuzner, M.L., Gveric, D., Strand, C., Loughlin, A.J., Paemen, L., Opdenakker, G., Newcombe, J. (1996) The expression of tissue-type plasminogen activator, matrix metalloproteases and endogenous inhibitors in the central nervous system in multiple sclerosis: comparison of stages in lesion evolution. *J. Neuropathol. Exp. Neurol.* **55**, 1194– 1204.
31. Belaaouaj, A., Shipley, J.M., Kobayashi, D.K., Zimonjic, D.B., Popescu, G.A., Silverman, G.A., Shapiro, S.D. (1995) Human macrophage metalloelastase. Genomic organization, chromosomal location, gene linkage, and tissue-specific expression. *J. Biol. Chem.* **270**, 14568–14575.
32. Lanone, S., Zheng, T., Zhu, Z., Liu, W., Lee, C.G., Ma, B., Chen, Q., Homer, R.J., Wang, J., Rabach, L.A., Rabach, M.E., Shipley, J.M., Shapiro, S.D., Senior, R.M., Elias, J.A. (2002) Overlapping and enzyme-specific contributions of matrix metalloproteinases-9 and -12 in IL-13-induced inflammation and remodeling. *J Clin Invest* **110**, 463-474.
33. Abboud, R.T., Ofulue, A.F., Sansores, R.H., Muller, N.L. (1998) Relationship of alveolar macrophage plasminogen activator and elastase activities to lung function and CT evidence of emphysema. *Chest* **113**, 1257–1263.
34. Barnes, P.J. (1998) Chronic obstructive pulmonary disease: new opportunities for drug development. *Trends Pharmacol Sci* **19**, 415–423.
35. MacNee, W. (2001) Oxidants/antioxidants and chronic obstructive pulmonary disease: pathogenesis to therapy, *Novartis Found. Symp.* **234** 169–185.

36. Hanrahan, J.P., Sherman, C.B., Bresnitz, E.A., *et al*. (1996) Cigarette smoking and health *Am. J. Respir. Crit. Care Med.* **153**, 861–865.
37. Ludwig, P.W., Schwartz, B.A., Hoidal, J.R., Niewoehner, D.E. (1985) Cigarette smoking causes accumulation of polymorphonuclear leukocytes in alveolar septum. *Am Rev Respir Dis* **131**, 828-830.
38. Eidelman, D., Saetta, M.P., Ghezzo, H., Wang, N.S., Hoidal, J.R., King, M., Cosio, M.G. (1990) Cellularity of the alveolar walls in smokers and its relation to alveolar destruction. Functional implications. *Am Rev Respir Dis* **141**, 1547-1552.
39. Finkelstein. R., Fraser, R.S., Ghezzo, H., Cosio, M.G. (1995) Alveolar inflammation and its relation to emphysema in smokers. *Am J Respir Crit Care Med* **152**, 1666-1672.
40. Tetley, T.D. (2002) Macrophages and the pathogenesis of COPD. *Chest* **121**, 156S-159S.
41. Hautamaki, R.D., Kobayashi, D.K., Senior, R.M., Shapiro, S.D. (1997) Requirement for macrophage elastase for cigarette smoke induced emphysema in mice. *Science* **277**, 2002-2004.
42. Thompson, R.W., Holmes, D.R., Mertens, R.A., Liao, S., Botney, M.D., Mecham, R.P., Welgus, H.G., Parks, W.C. (1995) Production and localization of 92-kD gelatinase in abdominal aortic aneurysms: an elastolytic metalloproteinase expressed by aneurysm-infiltrating macrophages. *J Clin Invest*. **96**, 318-326.
43. Curci, J.A., Liao, S., Huffman, M.D., Shapiro, S.D., Thompson, R.W. (1998) Expression and localization of macrophage elastase (matrix metaloproteinase-12) in abdominal aortic aneurysms. *J Clin Invest.* **102**, 1900-1910.
44. Gately, S., Twardowski, S.P., Stack, M.S., Patrick, M., Boggio, L., Cundiff, D.L., Schnaper, H.W., Madison, L., Volpert, O., Bouck, N., Enghild, J., Kwaan, H.C., Soff, G.A. (1996) Human prostate carcinoma cells express enzymatic activity that converts human plasminogen to the angiogenesisis inhibitor, angiostatin. *Cancer Res*. **56**, 4887-4890.
45. Stathakis, P., Fitzgerald, M., Mathias, L.J., Chesterman, C.N., Hogg, P.J. (1997) Generation of angiostatin by reduction and proteolysis of plasmin: catalysis by a plasmin reductase secreted by cultured cells. *J Biol Chem*. **272**, 20641-20645.

46. Patterson, B.C., Sang, Q.A. (1997) Angiostatin-converting enzyme activities of human matrilyisin (MMP-7) and gelatinase B/type IV collagenase (MMP-9). *J Biol Chem*. **272**, 28823-28825.
47. Cao, Y., Chen, A., An SSA, Ji R.W., Davidson, D., Llinas, M. (1997) Kringle 5 of plasminogen is a novel inhibitor of endothelial cell growth. *J Biol Chem*. **272**, 22924-22928.
48. Dong, Z., Kumar, R., Yang, X., Fidler, I.J. (1997) Macrophage-derived metalloelastase is responsible for the generation of angiostatin in Lewis lung carcinoma. *Cell*. **88**, 801-810.
49. Anderson, I.C., Shipp, M.A., Docherty, A.J.P., Teicher, B.A. (1996) Combination therapy including a gelatinase inhibitor and cytotoxic agent reduces local invasion and metastasis of murine Lewis lung carcinoma. *Cancer Res*. **56**, 710-715.
50. Kerkela, E., Ala-Aho, R., Jeskanen, L., *et al*. (2000) Expression of human macrophage metalloelastase (MMP-12) by tumor cells in skin cancer. *J Invest Dermatol* **114**, 1113–1119.
51. Kachra, Z., Beaulieu, E., Delbecchi, L., *et al*. (1999) Expression of matrix metalloproteinases and their inhibitors in human brain tumors. *Clin Exp Metastasis* **17**, 555–566.
52. Rivas, M.J., Arii, S., Furutani, M., *et al*. (1998) Expression of human macrophage metalloelastase gene in hepatocellular carcinoma: correlation with angiostatin generation and its clinical significance. *Hepatology* **28**, 986–993.
53. Gorrin-Rivas, M.J., Arii, S., Mori, A., *et al*. (2000) Implications of human macrophage metalloelastase and vascular endothelial growth factor gene expression in angiogenesis of hepatocellular carcinoma. *Ann Surg* **231**, 67–73.
54. O'Reilly, M.S., Holmgren, L., Shing, Y., *et al*. (1994) Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell* **79**,315–328.
55. Cornelius, L.A., Nehring, L.C., Harding, E., *et al*. (1998) Matrix metalloproteinases generate angiostatin: effects on neovascularization. *J Immunol* **161**, 6845–6852.

3. Methodological Aspects

3.1 Structural Genomics/Proteomics approach

The essence of structural genomics is to start from the gene sequence, produce the protein and determine its three-dimensional structure. The challenge, once the structure is determined, is to extract useful biological information about the biochemical and biological role of the protein in the organism. This is a complete reversal of the classical structural biology paradigm, where a protein structure is determined to understand how it performs its known biological function at the molecular level and its implications for drug design and target discovery. It is first useful to summarize what information can and cannot be derived from structural data (Fig. 9). The structure reveals the overall organization of the protein chain in three dimensions. From this we can identify the residues that are buried in the core or exposed to solvent on the protein surface, the shape and molecular composition of the surface, and the relative juxtaposition of individual groups. It also reveals the quaternary structure of the protein in the crystal environment or in solution at high concentration. Protein–ligand complexes are perhaps the most useful for functional information, since they reveal the nature of the ligand, where it is bound in the protein and, if the protein is an enzyme, the disposition of residues in the active site, from which a catalytic mechanism may be postulated. Before considering how to extract functional information, it is important to emphasize that structural data usually only carry information about the biochemical function of the protein. Its biological role in the cell or organism is much more complex and additional experimental information is needed in order to elucidate this. How then can structural information are used to infer functional knowledge? Comparison of the protein fold [(1)] or structural motifs within a protein to the structural databases may reveal similarities, from which biochemical and biological functional information may be inferred [(2)]. This knowledge-based approach is by far the most powerful method for predicting function from structure. All new structures are now scanned against the databases of known structures to find such relatives. The impact of structural genomics on target discovery and rational drug design is not clear. For target discovery, the hope is that the structures will reveal functions of proteins of relevance for therapy, thus providing new targets. The structures will also be valuable in deciding whether the protein is a suitable target for small molecule inhibitors (that is, ‘drugable’), through analysis of its binding site character.

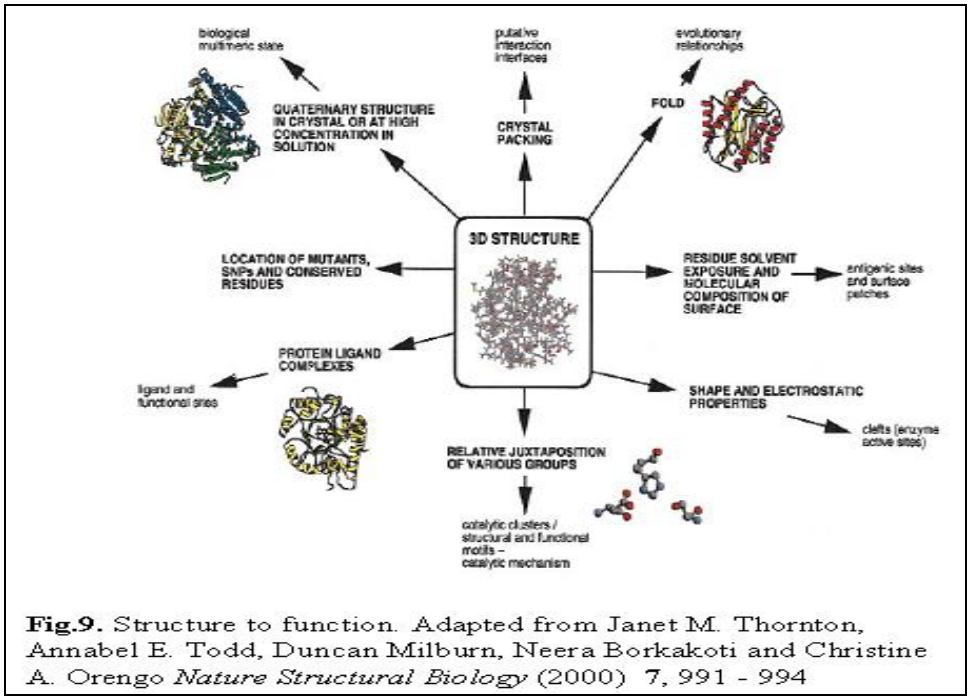


Fig.9. Structure to function. Adapted from Janet M. Thornton, Annabel E. Todd, Duncan Milburn, Neera Borkakoti and Christine A. Orengo *Nature Structural Biology* (2000) 7, 991 - 994

Knowledge of the distribution of proteins between species will be important auxiliary information to identify species-specific targets and to try to avoid toxicity problems. The structures may also prove useful in understanding and ultimately predicting cross-reactions of a small molecule with non-targeted proteins in the same organism, which can lead to multiple side effects. This will help to prioritize candidate targets, streamline the drug development pipeline, and hopefully reduce the number of failures downstream. For structure-based drug design, most chemists would argue that they need very high-resolution structures, preferably with multiple ligands. Here, although the high-throughput methods developed in the structural genomics projects will facilitate rapid structure determination, the need for high-resolution data is unlikely to change. However as more structures become available, it will gradually become the 'norm' for the development process for a new drug to incorporate knowledge of its target structure. Even today many of the newest drugs have been designed using a combination of combinatorial chemistry and intelligent structure-based design to bias libraries or to refine the lead molecule. Successful structure-based drug design *in silico* is still a goal to work toward, but as more structural data become available we can expect theoretical approaches to improve considerably.

3.2 Domain definition

As discussed above MMPs are structurally multi domains proteases composed minimally of an N-terminal signal sequence followed by the prodomain, catalytic domain and a C-terminal Hemopexin-like domain as in their simplest structure. In the current work various domains of MMP12 were cloned and expressed. There are various preliminary analyses needed in order to the best construct in terms of yield, solubility and purification.

Different constructs for MMP12 Catalytic (Cat) domain (Wild type and mutants), Hemopexin-like (Hpx) domain and Catalytic-Hemopexin-like (CatHpx) domain (Wild type and mutants) were cloned during the course of work. The different domains of the protein were defined and characterized by the use of available domain architectural bioinformatics tools such as Simple Modular Architectural Research Tool (SMART) [(3)] (http://smart.embl-heidelberg.de/) and/or Pfam (http://pfam.sanger.ac.uk/). On the basis of the results from the domain architectural tools, various constructs for domains were cloned and standardized for the expression in the prokaryotic host system (*Escherichia coli*, as in the current work).

3.3 Gene Cloning

Structural genomics consists in the determination of the three dimensional structure of all proteins of a given organism, by experimental methods such as X-ray crystallography, NMR spectroscopy or computational approaches such as homology modeling [(4)]. This raises new challenges in structural bioinformatics, i.e. determining protein function from its 3D structure. Structural genomics emphasizes high throughput determination of protein structures. Thus, the need of high throughput protein preparation is one of the challenging and emerging aspects of structural genomics or structural proteomics [(5)]. The ever fasting developments of techniques for cloning, expression, purification of proteins are thus the backbone of structural genomics or proteomics project. Molecular cloning refers to the procedure of isolating a defined DNA sequence and obtaining multiple copies of it *in vivo*. Cloning is frequently employed to amplify DNA fragments containing genes, but it can be used to amplify any DNA sequence [(6)].

The strategy used in gene cloning involves:

- the gene of interest is amplified by PCR.

- ❖ the PCR product is cloned into a specific cloning/expression vector and amplified.
- ❖ the sequence of a positive clone is checked by sequence analysis.
- ❖ the positive clone is transformed into a variety of expression systems.

Cloning can be done by any of the method:

1. Classical method using restriction enzyme
2. TA Cloning and TOPO TA Cloning
3. Recombination Cloning systems
    - Gateway Cloning Technology (Invitrogen)
    - Creator (BD Clontech)

In the present work, the classical cloning method (Fig. 10), which involves the use of restriction enzymes, has been used for cloning of different constructs of MMP12. Restriction enzymes (restriction endonucleases) are molecular scissors that cuts DNA at (or close to) specific recognition sites [7]. Two types of restriction enzymes exist that differ in the way they cut the target DNA:

- **Blunt end cutters.** These enzymes cut both strand of the target DNA at the same spot creating blunt ends.
- **Sticky end cutters.** These enzymes cut both strand of the target DNA at different spots creating 3'- or 5'-overhangs of 1 to 4 nucleotides (so-called sticky ends).

To be able to clone a DNA insert into a cloning or expression vector, both insert and vector have to be treated with two restriction enzymes that create compatible ends [8, 9]. The next step is the ligation of the insert into the linearized vector. This involves the formation of phosphodiester bonds between adjacent 5'-phosphate and 3'-hydroxyl residues, which can be catalyzed by ligase:

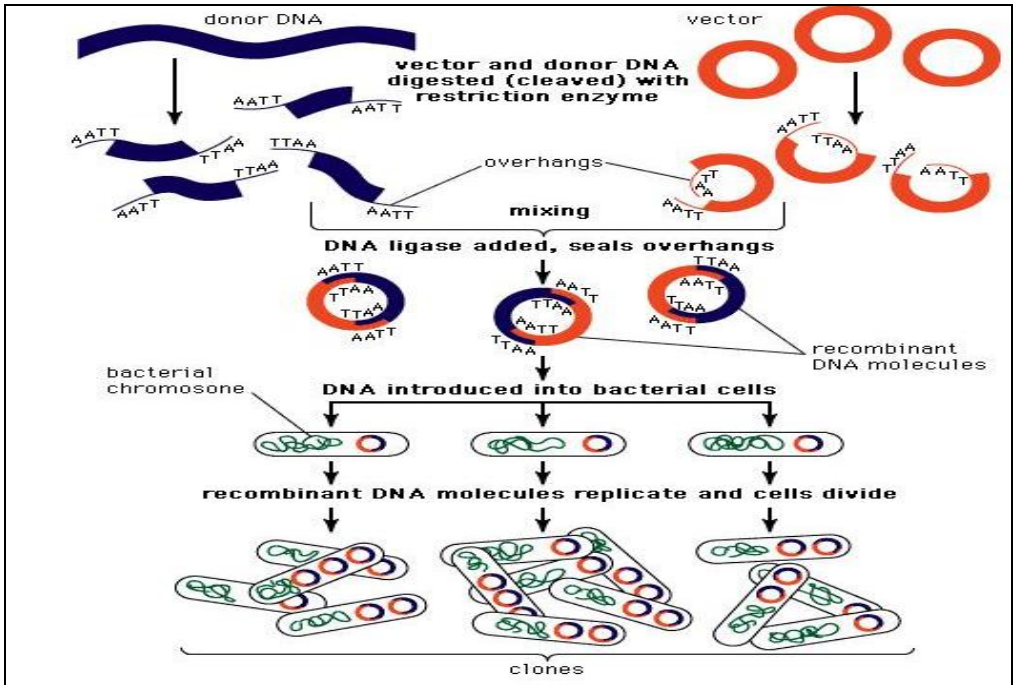


Fig.12 Steps involved in Molecular Cloning. Adapted from www.britannica.com

In the present work different domains of MMP12 (Cat, Hpx, and Cat-Hpx) were cloned using the above mentioned cloning strategy. *Nde*I and *Xho*I restriction enzymes (New England Biolabs) were used for all the constructs in pET21a vector (Novagen) using T4 DNA Ligase for ligation. All the constructs (Fig. 11) designed for the protein to be expressed as native means without any kind of fusion tag to enhance the solubility or ease of purification.

MMP12 Cat domain

MGPVWRKHYITYRINNYTPDMNREDVDYAIRKAFQVWSNVTPLKFSKINTGMADILVVFARGAHGDFHAFDGKGGILAHAFG
PGSGIGGDAHFDEDEFWTTHSGGTNLFLTAVHEIGHSLGLGHSSDPKAVMFPTYKYVDINTFRLSADDIRGIQSLYG

MMP12 Hpx domain

MEPALCDPNLSFDAVTTVGNKIFFFKDRFFWLKVSERPKTSVNLISSLWPTLPSGIEAAYEIEARNQVFLFKDDKYWLISNL
RPEPNYPKSIHSFGFPNFVKKIDAAVFNPRFYRTYFFVDNQYWRYDERRQMMDPGYPKLITKNFQGIGPKIDAVFYSKNKYY
YFFQGSNQFEYDFLLQRITKTLKSNSWFGC

MMP12 Cat-Hpx domain

MGPVWRKHYITYRINNYTPDMNREDVDYAIRKAFQVWSNVTPLKFSKINTGMADILVVFARGAHGDFHAFDGKGGILAHAFG
PGSGIGGDAHFDEDEFWTTHSGGTNLFLTAVHEIGHSLGLGHSSDPKAVMFPTYKYVDINTFRLSADDIRGIQSLYGDPKEN
QRLPNPDNSEPALCDPNLSFDAVTTVGNKIFFFKDRFFWLKVSERPKTSVNLISSLWPTLPSGIEAAYEIEARNQVFLFKDD
KYWLISNLRPEPNYPKSIHSFGFPNFVKKIDAAVFNPRFYRTYFFVDNQYWRYDERRQMMDPGYPKLITKNFQGIGPKIDAV
FYSKNKYYYFFQGSNQFEYDFLLQRITKTLKSNSWFGC

Fig11. Amino acid sequences for construct of various domain of MMP12. Initial Methionine is color coded and inserted during cloning.

The cloning vector with the inserted gene then transformed and amplified in *E.coli* DH5α strain and the positive clones were selected by streaking the transformed cells on the

selective antibiotic media plate (Ampicillin in the present case). The antibiotic resistant colonies suppose to have the gene of interest incorporated within was confirmed by PCR screening and checked on 1.5% Agarose Gel and further by sequence analysis (at PRIMM Biotech SRL, MILAN). The positive clone after sequencing results were transformed in various expression systems available in the laboratory for protein expression.

3.4 Site-directed mutation

MMP12 and in general almost all MMPs in their active form are prone to the autolysis or self-digestion. This makes it an unsuitable protein to work with for a longer duration of time Mutation of residue at active site within the catalytic domain without disturbance in the structure conformation is the best solution to over come the autolysis problem. Bioinformatics studies and models generated using bioinformatics tools as well as from literature survey it is discovered that mutation of **Glutamic Acid** to **Alanine** at **position 219 (E219A)** in the active site results in decreased activity to several magnitude though the structural conformation fully retained. Further bioinformatics studies shows that mutation of **Phenylalanine** to **Aspartic Acid** at **position 171 (F171D),** a site presented far upstream to the active site, results in better solubility of MMP12. The development of fast and easy site-directed mutagenesis technique helps in achieving the mutation.

Site-directed mutagenesis, also known as site-specific or oligonucleotides-directed mutagenesis is the technique of molecular biology which is used to make mutation at a defined site in the DNA molecule [(10)]. Discovered lately in 1970s, the method gained popularity in the recombinant DNA technology because of its numerous advantages over the classical random mutagenesis technique [(11)]. The method is used to make point mutation, switch amino acid, insertion, duplication, deletion of single or multiple amino acids.

The method implies the use of PCR with oligonucleotides "primers" that contain the desired mutation. As the primers are the ends of newly-synthesized strands, there should be a mismatch during the first cycle in binding the template DNA strand, after that first round, the primer-based strand (containing the mutation) would be at equal concentration to the original template. After successive cycles, it would exponentially grow, and after

25, would outnumber the original, unmutated strand in the region of 8 millions to 1, resulting in a nearly homogeneous solution of mutated amplified fragments.

Despite the fact that the PCR product is overwhelmingly composed of mutated plasmid, the template DNA is typically eliminated by enzymatic digestion with *Dpn*I, a restriction enzyme which cleaves only Dam methylated DNA [12]. The template, which is derived from an alkaline lysis plasmid preparation and therefore is methylated, is destroyed in this step, but the mutated plasmid is preserved because it was generated *in vitro* and is unmethylated as a result (Fig. 12).

The mutagenic oligonucleotide primers must be designed individually to incorporate the desired point mutation or degenerate codon. The following considerations should be made when designing mutagenic primers:

- All of the primers used for simultaneous mutagenesis must anneal to the same strand of the template plasmid. In most cases, primers binding to either strand will be incorporated into mutant plasmids with equal efficiency.
- Primers may be designed to bind to adjacent sequences or to well separate regions on the same strand of the template plasmid.
- Primers should be between 25 and 45 bases in length, with a melting temperature (T_m) of ≥75°C. Primers longer than 45 bases may be used, but using longer primers increases the likelihood of secondary structure formation, which may affect the efficiency of the mutagenesis reaction. Optimum primer sets for simultaneous mutagenesis should have similar melting temperatures.

The following formula is commonly used for estimating the T_m of primers:

$$T_m = 81.5 + 0.41(\%GC) - 675/N - \%\ \text{mismatch}$$

• *N* is the primer length in bases

• values for **%GC** and **% mismatch** are whole numbers

During the course of work single mutant **F171D,** single mutant **E219A,** double mutant **F171D, E219A** for catalytic domain as well as single mutant **F171D,** single mutant **E219A,** double mutant **F171D, E219A**, double mutant **F171D, N268D,** triple mutant **F171D, E219A, N268D** for Cat-Hpx domains of MMP12 has been designed as per above

mentioned strategy by either point mutation or amino acid switch using Stratagene's multi site-directed mutagenesis kit (Catalog #200514 and #200515).

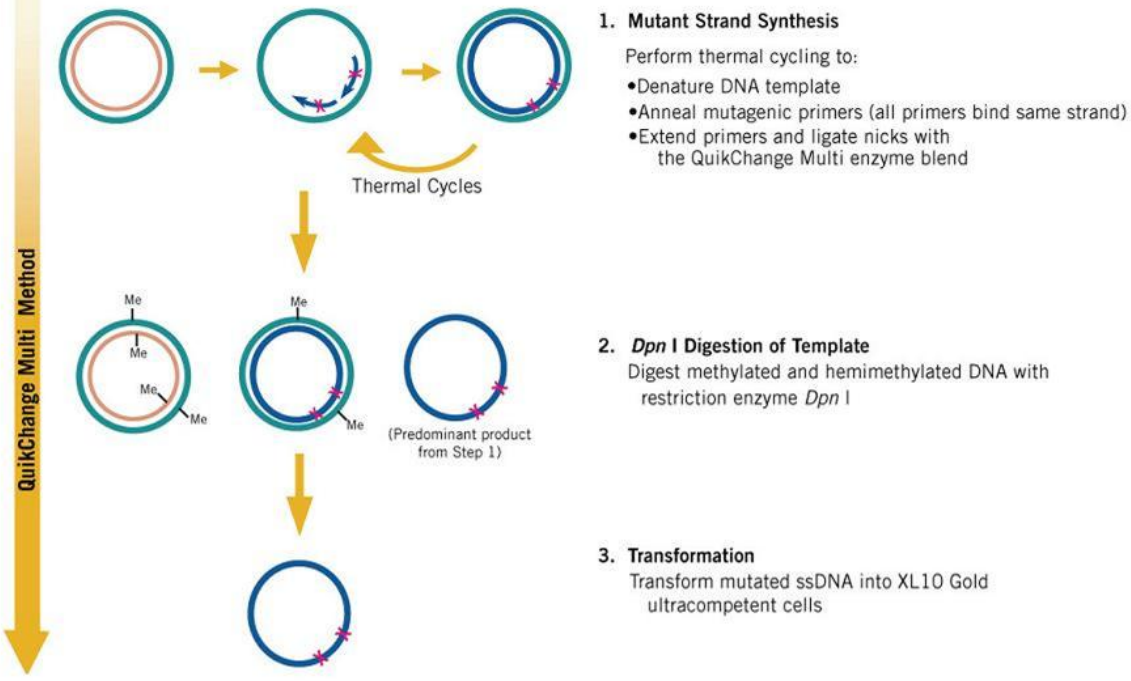


Fig12. Overview of Quick Multi Site-directed mutagenesis method. Adapted from Stratagene's Quick Multi Site-Directed Mutagenesis kit manual

3.5 Protein Expression

Protein overexpression the term given to the condition in which the protein encoded by a gene is expressed in a large amount. This can be achieved by increasing the copy number of the gene responsible for the production of particular protein or by increasing the binding strength of the promoter region [13]. Bacteria are good expression system for the overexpression of human recombinant proteins which required in large amount for the biophysical characterization using NMR or X-ray Crystallography or in a broad term in structural genomics. Overexpression is the state where about 30% of the total protein population is the desired protein.

Once the positive clone obtained, there are some preliminary analysis in order to standardize the expression conditions. The clones are transformed in a variety of expression host systems for comparisons of better expression. A small trial/test expression culture is usually perform in order to optimize the growth conditions which

includes agitation rate, temperature, inducer concentration, culture duration, etc. Protein from the small scale cultures is useful in optimizing the purification protocols. As the small scale culture got standardized the next step is to scale up the cultures using the refined conditions in shaker flask condition or in fermentors.

Here are the conditions optimized for overexpression of different domains of MMP12 in bacterial expression systems during the course of work:

MMP12 Catalytic domain (Gly106-Gly263):

The Cat domain (Wild-type, F171D single mutant, E219A single mutant, F171D, E219A double mutant) found to have optimal expression in *E.coli* BL21 (DE3) expression system. The cells were grown in LB medium containing 50 μg/ml of Ampicillin in a shaker flask at 37° C with agitation rate 180 rpm. Protein expression was induced with 0.5 mM IPTG at an OD_{600} = 0.6, and cell growth was continued for a further 5 h. For ^{13}C and ^{15}N labeled protein the amount of ^{13}C-enriched Glucose and ^{15}N-enriched Ammonium Chloride (Cambridge Isotopes Laboratories) were optimized to 2.5gm/l and 1.2gm/l of minimal media.

MMP12 Hemopexin-like domain (Glu278-Cys470):

The expression vector encoding the Hpx domain was transformed in *E.coli* BL21 (DE3) cells. Bacterial cells were grown in LB medium containing 50 μg/ml of Ampicillin in a shaker flask at 37° C with agitation rate 180 rpm. Protein expression was induced with 0.5 mM IPTG at an OD_{600} = 0.6, and cell growth was continued for overnight. Labeled culture in minimal media was done as metioned above for catalytic domain.

MMP12 Cat-Hpx domain (Gly106-Cys470):

Wild type Cat-Hpx as well as other 6 mutants for Cat-Hpx domain for optimal expression were transformed in *E.coli* BL21 (DE3) Codon Plus RIPL strain. The cells were grown at 37° C, 180 rpm in shaker flasks in LB medium containing 34 μg/ml of Chloramphenicol and 50 μg/ml of Ampicillin. Induction of protein carried out with 0.5 mM IPTG at an OD_{600} = 0.6, and cell growth was continued for overnight. Labeled culture in minimal media was done as mentioned above for catalytic domain.

3.6 Protein Purification

The strategy of purification depends mainly upon the location of the expressed protein within the host, indeed, the protein can be transported in the periplasmic space or expressed like a soluble or insoluble (inclusion bodies) protein within the cytoplasm [14]. In each case, the isolation was performed in a different way. All the purification involve several is chromatographic steps performed adjusting the parameters according to the different physical chemical and biological characteristics of the proteins. Ion exchange [15] (Fig. 13) and size exclusion [16] chromatography (Fig. 14) are commonly used to purify proteins expressed in their native states. Some tags, like MBP or GST, are specific affinity tag, which not only facilitate soluble expression but also increase the efficiency of protein purification. In some cases, other solubility tags have been combined with a simple His-tag, allowing the fusion partner to maintain its solubilizing functionality and also simultaneously as affinity tag.

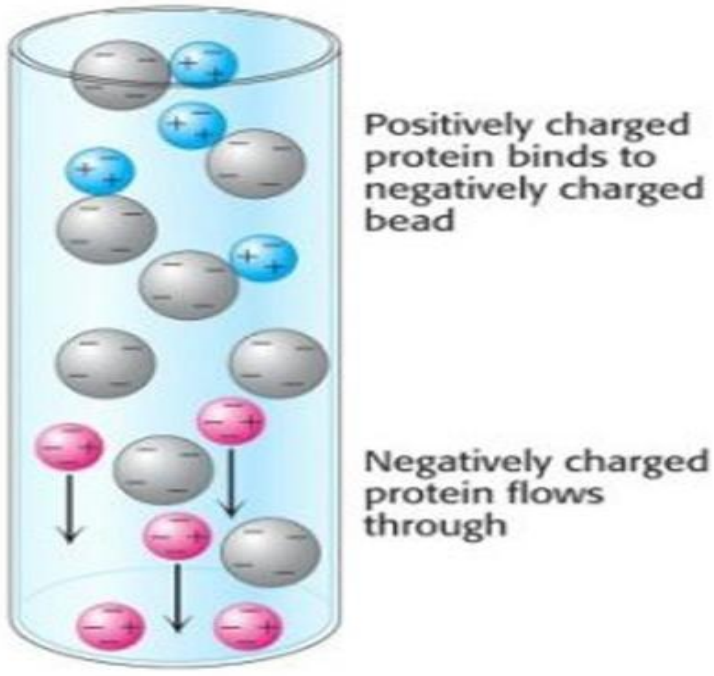


Fig.13. Principle of Ion exchange chromatography. Adapted from http://dbs.umt.edu/

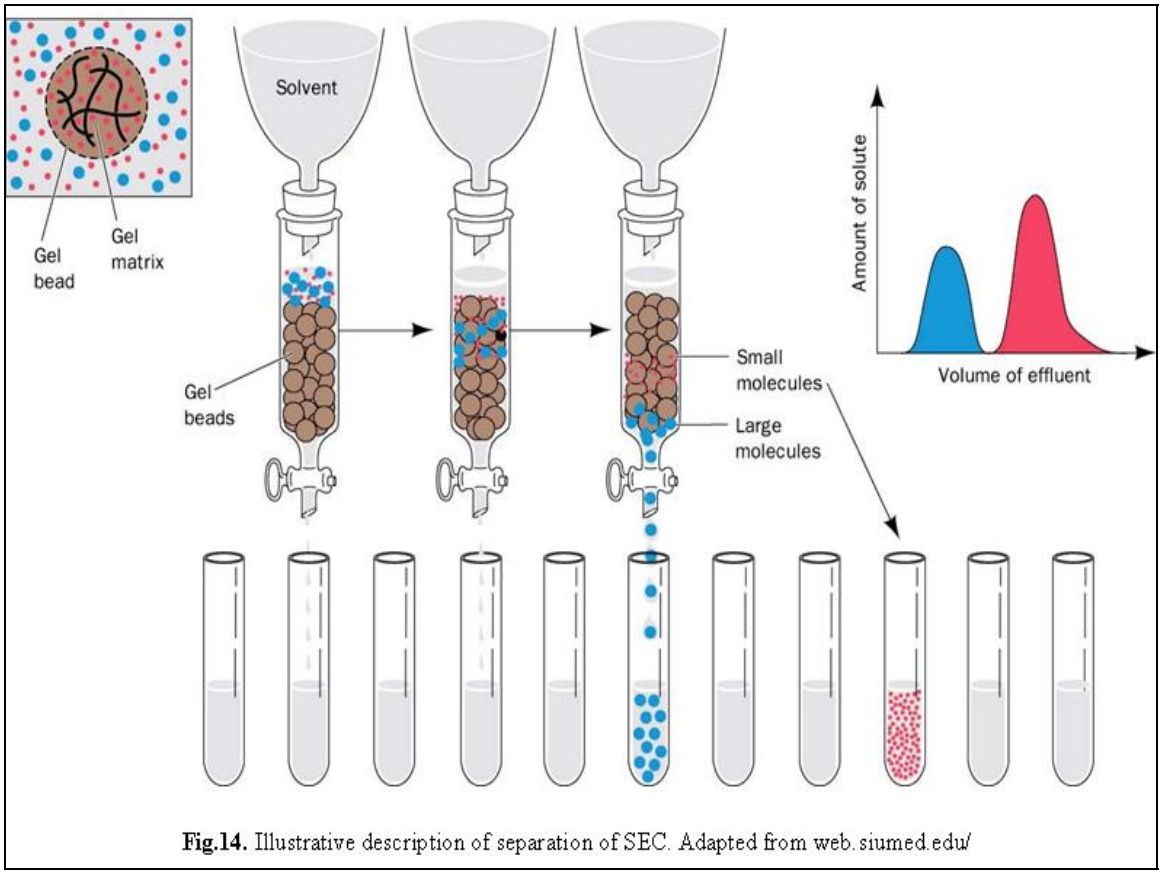


Fig.14. Illustrative description of separation of SEC. Adapted from web.siumed.edu/

For all domains of MMP12, irrespective of various fusion tags, the proteins were expressed as insoluble form in inclusion bodies. Thus to extract the protein from the inclusion bodies, protein was denatured using strong denaturing agent such as 8M Urea or 6M Guandinium Chloride. This result in an unfolded protein which needs to be refolds to its original conformation. Hence, refolding protocols were optimized. Here described are the protein purification methods for individual domains:

MMP12 Cat (Wild type and Mutants):

Cells were harvested by centrifugation, resuspended in 25% Sucrose, 50mM Tris-HCl (pH 8), 0.1M NaCl, 0.2M EDTA, 1mM DTT. 5-10 mg Lysozyme was added to the resulting suspension and stirred for 15-20 min. at 4°C. Buffer containing 2% Triton, 50mM Tris-HCl (pH 8), 0.1M NaCl, 0.2M EDTA, 1mM DTT was added and sonicated (7-8 cycles of 30 seconds) and centrifuged at 40000 rpm for 20 minutes at 4°C. The pellets were washed with 50mM Tris, 5mM EDTA, 1mM DTT (pH 8) and centrifuged. The resulting inclusion bodies were solublized in 8M Urea, 20mM Sodium Acetate (pH 5). The insoluble material was removed by centrifuging at 9000 rpm for 10 minutes. Protein molecular weight and purity was checked on 15% Gel by SDS-PAGE.

Protein was purified with Cation Exchange chromatography using HiTrap SP column using a gradient of NaCl up to 350mM (6M Urea, 20mM Sodium Acetate, pH 5). The purified protein was refolded by serial dilution method. The protein was diluted to a final concentration of 0.10mg/ml in 4M Urea, 20mM Tris-HCl, 10mM $CaCl_2$, 0.1mM $ZnCl_2$, 0.3M NaCl (pH 7.2), dialysed overnight against same buffer (4M Urea, 20mM Tris-HCl, pH 7.2, 10mM $CaCl_2$, 0.1mM $ZnCl_2$, 0.3M NaCl) successively followed by two steps with Buffer B (2M Urea, 20mM Tris pH 7.2, 10mM $CaCl_2$, 0.1mM $ZnCl_2$, 0.3M NaCl) for 8 hours, then finally 3 step of 8 hours each with Buffer C (20mM Tris pH 7.2, 10mM $CaCl_2$, 0.1mM $ZnCl_2$, 0.3M NaCl,). The protein was concentrated upto 5 ml and purified by size exclusion chromatography using High Load™ 16/60 Superdex™ 75ng (Amersham Biosciences) and eluted with 20mM Tris pH 7.2, 10mM $CaCl_2$, 0.1mM $ZnCl_2$, 0.3M NaCl and 0.2M AHA. The eluted fractions were checked for purity on 15% Gel by SDS-PAGE (Fig. 15) and elution fractions containing MMP-12-Cat were pooled and concentrated. The protein yield of well refolded purified protein was estimated about 100-120 mg/l for Wild type as well the mutants of catalytic domain (Table 1).

MMP12 Hpx & Cat-Hpx (Wild type and Mutants):

Overnight grown cells were harvested by centrifugation, resuspended in 25% Sucrose, 50mM Tris-HCl (pH 8), 0.1M NaCl, 0.2M EDTA, 1mM DTT. 5-10 mg Lysozyme was added to the resulting suspension and stirred for 15-20 min. at 4°C. Buffer containing 2% Triton, 50mM Tris-HCl (pH 8), 0.1M NaCl, 0.2M EDTA, 1mM DTT was added and sonicated (7-8 cycles of 30 seconds) and centrifuged at 40000 rpm for 20 minutes at 4°C. The pellets were resuspended in 6M Urea, 20mM Tris-HCl (pH 8) and centrifuged. The resulting inclusion bodies were solublized in 6M Gdn-HCl, 20mM Tris-HCl (pH 8), 10mM DTT, and 20mM Cystamine and stirred overnight at 4°C. The insoluble material was removed by centrifuging at 9000 rpm for 10 minutes. Protein molecular weight and purity was checked on 15% Gel by SDS-PAGE.

Refolding process was done by serial dilution method. The protein was diluted to a final concentration of 0.13mg/ml in 6M Gdn-HCl, 20mM Tris-HCl (pH 8), 1mM DTT, and 0.05% Brij-35 stirred at 4°C for 30-60 minutes and dialysed overnight against Buffer A (20mM Tris pH 7.2, 10mM $CaCl_2$, 0.1mM $ZnCl_2$, 0.15M NaCl, 5mM b-mercaptoethanol, 1mM 2-hydroxyethyl disulfide, 0.05% Brij-35) successively followed by Buffer B

(20mM Tris pH 7.2, 10mM $CaCl_2$, 0.1mM $ZnCl_2$, 0.15M NaCl, 1mM 2-hydroxyethyl disulfide) for 8 hours, Buffer C (20mM Tris pH 7.2, 10mM $CaCl_2$, 0.05mM $ZnCl_2$, 0.05M NaCl,) 5 hours and then finally 3 step of 5 hours each with buffer D (20mM Tris pH 7.2, 10mM $CaCl_2$). The refolded proteins was loaded on CM Sepharose Column (Amersham) (5ml X 2) equilibrated with 20mM Tris pH 7.2, 10mM $CaCl_2$ for the removal of Brij and the process was carried out at 4°C. The bounded protein was eluted by 20mM Tris pH 7.2, 10mM $CaCl_2$, 0.3M NaCl and 0.1M AHA. The protein was concentrated upto 5 ml and purified by size exclusion chromatography using High Load^{TM} 16/60 Superdex^{TM} 75ng (Amersham Biosciences) and eluted with 20mM Tris pH 7.2, 10mM $CaCl_2$, 0.3M NaCl and 0.2M AHA. The eluted fractions were checked for purity on 15% Gel by SDS-PAGE (Fig. 15) and elution fractions containing MMP-12-CatHpx were pooled and concentrated. The yield of well refolded and purified Hpx and Cat-Hpx (Wild type & mutants) domains were estimated about 60-80 mg/l and 80-100 mg/l respectively (Table 1).

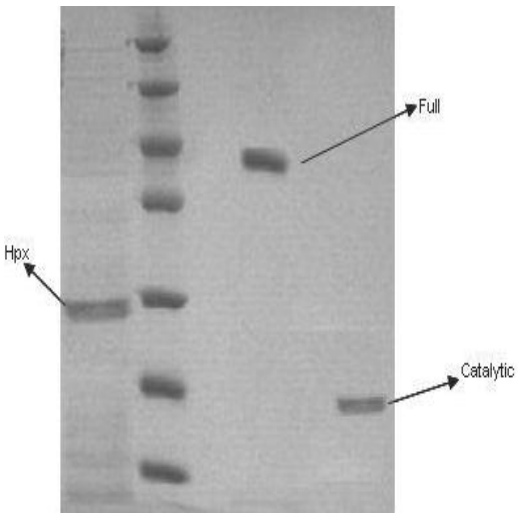


Fig.15. SDS PAGE for different domains of MMP12

MMP12		
DOMAIN	**FORM**	**YIELD (mg/liter)**
Catalytic	Wild Type	120-135
	F171D (Active)	
	E219A (Inactive)	
	F171D, E219A (Inactive)	
Hpx-like	Wild Type	60-80
Cat-Hpx	Wild Type	80-100
	F171D (Active)	
	E219A (Inactive)	
	F171D, E219A (Inactive)	
	F171D, N268D (Active)	
	F171D, E219A, N268D (Inactive)	

Table1. Different domains of MMP12 (Wild type and

3.7 Protein Stabilization

Most of the members of MMP family reported to be very active proteases and have the tendency of undergoing the process of autolysis or self degradation [16]. Thus, the problem of autolysis makes the class of protease an unstable protein for long term duration. MMP12 like most of its family member is very active and undergoes a process of self degradation in its active form resulting in individual catalytic as well as

hemopexin domain with an almost negligible amount of full length Cat-Hpx if kept for a long time. Though use of weak inhibitor like Acetohydroxamate results in a decreased activity but not very effective for long duration which is one important requirement of some NMR experiments as well as crystallization process. The problem is overcome by use of strong inhibitor for MMPs like NNGH but it not recommended for the interaction studies. Mutation at active sites with out change in structural conformation is one of the best options to work with MMPs. Though the mutation at active site (**E219A**) results in decrease in activity to several orders of magnitude but still the protease is not completely inactive and thus not suitable for substrate or ligand interaction studies. In order to overcome the problem, during the course of work protocol was standardized to have a stable protein for a longer duration. The troubleshoot lies in the replacement of the catalytic zinc with any other cadmium resulting in almost deceased protease activity without any affect on structural conformation or binding with substrate or ligand.

Cadmium, one of the most lethal metals for any form of life, placed in same group down to that of the zinc in the element periodic table. It is preferable over other metals to exchange the catalytic zinc ion in MMPs because of the following facts:

- It's in the same group to that of zinc in periodic table,
- It has the same divalent charge like to that of zinc,
- When ionized cadmium is almost same size to that of zinc,
- It doesn't exhibit catalytic activity.

Thus, taking in account the above mentioned criteria we optimized a protocol to exchange the catalytic zinc ion with cadmium ion. Again substitution of the catalytic zinc is not an easy method. Extensive dialysis of protein against 1mM cadmium for 3-4 weeks results in replacement of zinc with cadmium. Cadmium substitution was carried out by a slow dialysis process for 20-25 days. The concentrated protein (25-30µM) was dialyzed against buffer E (50mM HEPES, 0.2M AHA, 0.3M NaCl, pH 6.8 passed from Chelex-100 resin, 10mM $CaCl_2$ and 1mM $CdCl_2$) for 20-25 successive days (7ml of protein against 50ml buffer). After the extensive dialysis protein was centrifuge to remove any kind of precipitation and then concentrated up to 2 ml. The concentrated protein sample was washed 4-5 times with buffer F (20mM Tris-HCl, 0.2M AHA passed from Chelex-100 resin, 1mM $CdCl_2$) (2ml concentrated protein diluted up to 10 ml with buffer F and concentrated again using centricon) to remove the excess of Cadmium and to exchange

the buffer for the NMR sample. The cadmium substituted sample along with mutation at active site found to be stable for a longer duration at room temperature as well at 4° C.

3.8 Sample Preparation

NMR Sample. The NMR samples for various experiments with zinc or cadmium/cobalt substitution were prepared after purification by concentrating the protein and removing any precipitation resulting during concentration process. The final protein concentration for catalytic domain was 1 mM while 0.5-0.7 mM for Hpx and Cat-Hpx domain.

Crystallization. For MMP12–AHA adduct an aliquot of 2 µl of protein solution (20 mM Tris/10 mM $CaCl_2$/0.1 mM $ZnCl_2$/300 mM NaCl/200 mM AHA, pH 8) was mixed with 2 µl of reservoir buffer (0.1 M Tris·HCl/30% PEG 8000, pH 8). The final protein concentration was 1mM. Crystallization was carried out with the hanging drop vapor diffusion method at 20°C.

3.9 Biophysical Characterization

X-ray crystallography and NMR spectroscopy are the two main techniques that can provide structures of macromolecules at atomic resolution. Both techniques are well established and play a key role in structural biology as a basis for a detailed understanding of molecular functions. Their respective different advantages and disadvantages in terms of sample preparation and data collection and analysis make this it's complementary in structural proteomics. X-ray studies usually requires substantial investment of time in order to optimize the crystallization conditions and obtain a crystal with a good diffraction properties This process can take weeks or months, but once a well-diffracting (<2.5Å) crystal is obtained, the structure determination can proceed quite quickly. NMR Whereas X-ray crystallography requires single crystal; NMR measurements are carried out in solution under conditions that can be as close as possible to the physiological state. Some times even if crystal structures are available, additional data would be needed to determine the potential biological function of the protein. NMR is not only capable of solving protein structures to atomic resolution but also has the unique ability to accurately measure the dynamic properties of proteins and can also supply information on protein folding and on intra-, as well as, intermolecular interactions. Furthermore, the analysis through NMR spectroscopy easily allows the characterization under several, different experimental conditions, such as different ionic strength and pH. However, the two majors bottlenecks limiting the NMR in structural

biology are represent from the long time request for the data analysis and structural calculation and the drawback size limit of protein. The current is around 35KDa but recent advances in both hardware and experimental design promise to allow the study of much larger proteins [17].

Circular Dichroism (CD) Spectroscopy:

Circular Dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light which arise due to structural asymmetry. In addition, the difference in left and right handed absorbance A (l)-A (r) is very small (usually range of 0.0001) corresponding to an ellipticity of a few $1/100^{th}$ of a degree. The CD is a function of wavelength. CD spectra for distinct types of secondary structure present in peptides, proteins and nucleic acids are different. The analysis of CD spectra can therefore yield valuable information about secondary structure of biological macromolecules [18]. When this light passes through an optically active sample with a different absorbance A for the two components, the amplitude of the stronger absorbed component will smaller than that of the less absorbed component. The consequence is that a projection of the resulting amplitude now yields an ellipse instead of the usual line. The occurrence of ellipticity is called Circular Dichroism; it is not the same as optical rotation. Rotation of the polarization plane by a small angle α occurs when the phases for the 2 circular components become different, which requires a difference in the refractive index n. This effect is called circular birefringence. CD is measured as a quantity called *mean residue ellipticity*, whose units are degrees-cm^2/dmol. Chiral or asymmetric molecules produce a CD spectrum because they absorb left and right handed polarized light to different extents and thus are considered to be "optically active". Biological macromolecules such as proteins and DNA are compose of optically active elements and because they can adopt different types of three-dimensional structures, each type of molecule produces a distinct CD spectrum. Circular Dichroism spectroscopy is particularly good for:

- determining whether a protein is folded, and if so characterizing its secondary structure, tertiary structure, and the structural family to which it belongs,
- comparing the structures of a protein obtained from different sources or comparing structures for different mutants of the same protein,

- studying the conformational stability of a protein under stress-thermal stability, pH stability, and stability to denaturants and how this stability is altered by buffer composition or addition of stabilizers or excipients.
- Determining whether protein-protein interactions alter the conformation of protein. If there are any conformational changes, this will result in a spectrum which will differ from the sum of the individual components [19].

Secondary structure of a protein can be determined by CD spectroscopy in the "far-UV" spectral region (190-250nm). At these wavelengths the chromophore is the peptide bond, and the signal arises when it is located in a regular folded environment.
Alpha helix, beta-sheet and random coil structures each give rises to a characteristic shape and magnitude of CD spectrum. The approximate fraction of each secondary structure type that is present in any protein can thus be determined by analyzing its far-UV spectrum as a sum of fractional multiples of such reference spectra for each structural type. Like all spectroscopic techniques, the CD signal reflects an average of the entire molecular population. Thus, while CD can determine that a protein contains about 50% alpha-helix, it cannot determine which specific residues are involved in the alpha helical portion [20].

X-ray and protein crystallization:

X-ray crystallography enables to visualize protein structures at the atomic level and enhances our understanding of protein function [21]. Specifically it is possible to study how proteins interact with other molecules, how they undergo conformational changes, and how they perform catalysis in the case of enzymes. In order to see proteins in atomic detail we need to work with electro-magnetic radiation with a wavelength of around 0.1 nm or 1 Å [22].

The diffraction from a single molecule would be too weak to be measurable. So it is necessary to use an ordered three-dimensional array of molecules, in other words a crystal, to magnify the signal. Even a small protein crystal might contain a billion molecules. If the crystal is well ordered, then diffraction will be measurable at high angles or high resolution and a detailed structure should result. The X-rays are diffracted

by the electrons in the structure and consequently the result of an X-ray experiment is a 3-dimensional map showing the distribution of electrons in the structure [23].

X-ray studies usually require substantial investment of time in order to optimize the crystallization conditions and obtain a crystal with good diffraction properties. This process can take weeks or months, but once a well-diffracting (<2.5Å) crystal is obtained; the structure determination can proceed quite quickly. In order to crystallize a protein, the purified protein undergoes slow precipitation from an aqueous solution [24]. The production of good crystals is dependent upon a number of environmental factors because so much variation exists among proteins, with each individual requiring unique condition for successful crystallization. Especially protein purity and pH conditions are very important; for example different pH values can result in different packing orientations.

The two most commonly used methods for protein crystallization are both vapor diffusion techniques. These are known as the "hanging drop" and **"**sitting drop" methods [25]. Both entail a droplet containing purified protein, buffer and precipitant being allowed to equilibrate with a larger reservoir containing similar buffers and precipitants in higher concentrations. Initially, the droplet of protein solution contains an insufficient concentration of precipitant for crystallization, but as water vaporizes from the drop and transfers to the reservoir, the precipitant concentration increases to a level optimal for crystallization. Since the system is in equilibrium, these optimum conditions are maintained until the crystallization is complete.

Nuclear Magnetic Resonance (NMR) Spectroscopy:

Protein nuclear magnetic resonance spectroscopy (usually abbreviated protein NMR) is a field of structural biology in which NMR spectroscopy is used to obtain information about the structure and dynamics of proteins. Structure determination by NMR spectroscopy usually consists of several following phases, each using a separate set of highly specialized techniques. The sample is prepared, resonances are assigned, restraints are generated and a structure is calculated and validated.

Protein NMR utilizes multidimensional nuclear magnetic resonance experiments to obtain information about the protein. Ideally, each distinct nucleus in the molecule experiences a distinct chemical environment and thus has a distinct chemical shift by

which it can be recognized. However, in large molecules such as proteins the number of resonances can typically be several thousand and a one-dimensional spectrum inevitably has incidental overlaps. Therefore multidimensional experiments are performed which correlate the frequencies of distinct nuclei. The additional dimensions decrease the chance of overlap and have larger information content since they correlate signals from nuclei within a specific part of the molecule. Pulse sequences allow the experimenter to investigate and select specific types of connections between nuclei. Depending on the concentration of the sample, on the magnetic field of the spectrometer, and on the type of experiment, a single multidimensional nuclear magnetic resonance experiment on a protein sample may take hours or even several days to obtain suitable signal-to-noise ratio through signal averaging, and to allow for sufficient evolution of magnetization transfer through the various dimensions of the experiment. Other things being equal, higher-dimensional experiments will take longer than lower-dimensional experiments.

Typically the first experiment to be measured with an isotope-labeled protein is a 2D heteronuclear single quantum correlation (HSQC) spectrum where "heteronuclear" refers to nuclei other than ^{1}H. In theory the heteronuclear single quantum correlation has one peak for each H bound to a heteronucleus [26]. Thus in the ^{15}N-HSQC one signal is expected for each amino acid residue with the exception of proline which has no amide-hydrogen due to the cyclic nature of its backbone. The ^{15}N-HSQC is often referred to as the fingerprint of a protein because each protein has a unique pattern of signal positions.

With unlabelled protein the usual procedure is to record a set of two dimensional homonuclear nuclear magnetic resonance experiments through correlation spectroscopy (COSY), of which several types include conventional correlation spectroscopy, *total correlation* spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY)[27, 28]. A two-dimensional nuclear magnetic resonance experiment produces a two-dimensional spectrum. The units of both axes are chemical shifts. The COSY and TOCSY transfer magnetization through the chemical bonds between adjacent protons. The process of resonance assignment for a nitrogen-15 labeled sample is similar to the homonuclear case. No experiment can be performed that transfers magnetization between two spin systems through bonds either. The main difference is the ability to record nitrogen-15 edited three dimensional experiments: TOCSY-N HSQC and NOESY-N-HSQC. These experiments build onto the HSQC experiment, but have an additional

proton dimension. It can be visualised as each peak in the HSQC having the TOCSY or NOESY peaks stacked onto it. Thus if the TOCSY peak from an amide proton, H_N, has a cross peak to its alpha proton, H_{alpha}, at the coordinates (H_N, H_{alpha}) in the TOCSY spectrum, the corresponding peak would be at (H_N, H_{alpha},N) in the TOCSY-N-HSQC. Thus it is possible to resolve overlaps in the proton dimension, if the corresponding nitrogen has chemical shifts distinct from one another. When the protein is labeled with carbon-13 and nitrogen-15 it is possible to record an experiment that transfers magnetization over the peptide bond, and thus connect different spin systems through bonds. This is usually done using some of the following experiments, HNCO, HNCACO, HNCA, HNCOCA, HNCACB and CBCACONH [29, 30]. All six experiments consist of a HSQC plane expanded with a carbon dimension. In the HNCACO the spectrum contains peaks at the chemical shifts of the carbonyl carbons in the residue of the HSQC peak and the previous one in the sequence. The HNCO only contains the chemical shift from the previous residue, and it is thus possible to assign the carbonyl carbon shifts that correspond to each HSQC peak and the one previous to that one. Sequential assignment can then be undertaken by matching the shifts of each spin system's own and previous carbons. The HNCA and HNCOCA work similarly, just with the alpha carbons rather than the carbonyls, and the HNCACB and the CBCACONH contains both the alpha carbon and the beta carbon. Usually several of these experiments are required to resolve overlap in the carbon dimension. This procedure is usually less ambiguous than the NOESY based method, since it is based on through bond transfer. In the NOESY-based methods additional peaks that are close in space but not belonging to the sequential residues will appear confusing the assignment process. When the sequential assignment has been made it is usually possible to assign the side-chains using HCCH-TOCSY, which is basically a TOCSY experiment resolved in an additional carbon dimension.

REFERENCES:

1. Orengo, C.A. (1994) *Curr. Opin. Struct. Biol.* **4**, 429–440.
2. Holm, L. & Sander, C. (1999) *Nucleic Acids Res*. **27**, 244–247.
3. Jörg Schultz, Frank Milpetz, Peer Bork, Chris P. Ponting (1998) *Proc.Natl.Acad.Sci.* **95**, 5857-5864.
4. Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., Sali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annu Rev Biophys Biomol Struct* **29**, 291-325.
5. Liebler, D. C. (2002) Introduction to proteomics: tools for the new biology. Humana Press, Totowa, NJ.
6. Peter J. Russel (2005). *iGenetics: A Molecular Approach*. San Francisco, California, Pearson Education.
7. Roberts, R.J. (1976). "Restriction endonucleases". *CRC Crit. Rev. Biochem.* **4** (2), 123–64.
8. Arber, W., Linn, S. (1969). "DNA modification and restriction". *Annu. Rev. Biochem.* **38**, 467–500.
9. Primrose, Sandy B.; Old, R. W. (1994). *Principles of gene manipulation: an introduction to genetic engineering*. Oxford: Blackwell Scientific.
10. Zheng, L., Baumann, U., Reymond, J.L. (2004) *Nucleic Acids Res* Aug 10; **32** (14), e115
11. Ko, J.K., Ma, J. (2005) *Am J Physiol Cell Physiol* **288** (6), C1273-8.
12. Friedman, K. L., M. K. Raghuraman, W. L. Fangman, and B. J. Brewer (1995) *J. Cell Sci. Suppl.* **19**, 51-58.
13. Holz, C.; Hesse, O.; Bolotina, N.; Stahl, U.; Lang, C. (2002) *Protein Expression & Purification*. **25**, 372-378.
14. Qing, G.; Ma, L.C.; Khorchid, A.; Swapna, G.V.; Mal, T.K.; Takayama, M.M.; Xia, B.; Phadtare, S.; Ke H.; Acton T.; (2004) *Nat Biotechnol* **22,** 877–882.
15. Helfferich, F. (1962) *Ion Exchange* McGraw Hill, New York.
16. Eisenstein, M. (2006) *Nature Methods* **3**(5): 410.
17. Wuthrich, K. (1998) *Nat. Struct. Biol.* **5**, 492–495.
18. Canter, C.R., Schimmel, P.R. (1980) Biophysical Chemistry Volume 2, Chapter 10 W.H. Freeman & Company, NY.

19. Alder, A.J., Greenfield, N.J., Fasman, G.D. (1973) *Methods in Enzymology* **27**, 675.
20. Johnson, W.C. (1990) *Proteins* **7**, 205-214.
21. Geerlof, A., Brown, J., Coutard, B., Egloff, M.P., Enguita, F.J., Fogg, M.J., Gilbert, R.J., Groves, M.R., Haouz, A., Nettleship, J.E., Nordlund, P., Owens, R.J., Ruff, M., Sainsbury, S., Svergun, D.I., Wilmanns, M. (2006) *Acta Crystallogr. D Biol. Crystallogr.* **62** (Pt 10): 1125–36.
22. Scapin, G. (2006) *Current Pharmaceutical Design* **12**, 2087-2097.
23. Rhodes, C.J., Jacobs, R.L. (1993) *Journal of Physics-Condensed Matter* **5** 5649-5662.
24. Branden, Tooze, (1999) *Introduction to Protein Structure*, New York pp. 374-376.
25. D. E. McRee, D.E. (1993) *Practical Protein Crystallography*, Academic Press Inc., San Diego.
26. Cavanagh, J., Fairbrother, W.J., Palmer, A.G., Skelton, N.J. (1995) Protein NMR Spectroscopy: Principles and Practice Academic Press.
27. Wuthrich K. (1990) *J Biol Chem.* **265** (36), 22059-62.
28. Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley New York.
29. Guntert, P. (2004) *Methods Mol Biol.* **278**, 353-78.
30. Liu, G., Shen, Y., Atreya, H.S., Parish, D., Shao, Y., Sukumaran, D.K., Xiao, R., Yee, A., Lemak, A., Bhattacharya, A., Acton, T.A., Arrowsmith, C.H., Montelione, G.T., Szyperski, T. (2005) *Proc Natl Acad Sci* **102**(30), 10487-92.

4. Evidence of reciprocal reorientation of the catalytic and hemopexin-like domains of full length MMP-12[1]

Ivano Bertini*, Vito Calderone, Marco Fragai, **Rahul Jaiswal,** Claudio Luchinat, Maxime Melikian, Estradifois Mylonas, Dmitris I. Svergun

Journal of American Chemical Society, (2008) **130** (22), 7011-21

---

[1], I.B., C.L. & M.F. designed the research, R.J. done all the Molecular Biology work and sample preparation, V.C. done the X-ray crystallography experiment, M.M. & M.F. done the NMR experiments, E.M. & D.I.S done the SAXS experiments.

J|A|C|S
ARTICLES

Published on Web 05/09/2008

Evidence of Reciprocal Reorientation of the Catalytic and Hemopexin-Like Domains of Full-Length MMP-12

Ivano Bertini,*[,†,‡] Vito Calderone,[†] Marco Fragai,[†,§] Rahul Jaiswal,[†] Claudio Luchinat,[†,§] Maxime Melikian,[†] Efstratios Mylonas,[||] and Dmitri I. Svergun[||,⊥]

Magnetic Resonance Center (CERM), University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy, Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Italy, Department of Agricultural Biotechnology, University of Florence, Via Maragliano, 75−77, 50144 Florence, Italy, European Molecular Biology Laboratory, Hamburg Outstation, Notkestrasse 85, 22603 Hamburg, Germany, and Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, 117333 Moscow, Russia

Received November 20, 2007; E-mail: ivanobertini@cerm.unifi.it

Abstract: The proteolytic activity of matrix metalloproteinases toward extracellular matrix components (ECM), cytokines, chemokines, and membrane receptors is crucial for several homeostatic and pathological processes. Active MMPs are a family of single-chain enzymes (23 family members in the human genome), most of which constituted by a catalytic domain and by a hemopexin-like domain connected by a linker. The X-ray structures of MMP-1 and MMP-2 suggest a conserved and well-defined spatial relationship between the two domains. Here we present structural data for MMP-12, suitably stabilized against self-hydrolysis, both in solution (NMR and SAXS) and in the solid state (X-ray), showing that the hemopexin-like and the catalytic domains experience conformational freedom with respect to each other on a time scale shorter than 10^{-8} s. Hints on the probable conformations are also obtained. This experimental finding opens new perspectives for the often hypothesized active role of the hemopexin-like domain in the enzymatic activity of MMPs.

Introduction

Matrix metalloproteinases (MMP) are an important family of 23 proteins which are involved in a number of extracellular processes[1–3] including the degradation of the extracellular matrix.[4] The latter is constituted by structural proteins such as collagen and elastin, by proteoglycans, and by adhesive proteins such as fibronectin, laminin, and tenascin.[5] MMPs are single-chain enzymes secreted by cells as inactive proenzymes. The active form is liberated outside the cell by the cleavage of the prodomain by other proteases, including MMPs themselves,[6,7] thereby implying a complex regulation mechanism which also involves other proteins such as tissue inhibitors of MMPs (called TIMPs).[8,9]

All active MMPs but MMP-7 are constituted by two domains, a catalytic (CAT) and a hemopexin-like (HPX) domain. The CAT and HPX domains are connected by a linker whose length varies from 14 to 68 AA.[10,11] For many MMPs the linker is relatively short (14−23 AA) whereas for MMP-9 and MMP-15, at the other extreme, the intervening residues between the CAT and HPX domains (68 and 63 AA, respectively) constitute a further, highly glycosylated, domain termed OG domain.[12,13] The CAT domain alone bears full proteolytic activity toward a range of peptides and proteins.[14–17] However, efficient proteolysis of, for instance, triple helical collagen requires the full-

[†] CERM, University of Florence.
[‡] Department of Chemistry, University of Florence.
[§] Department of Agricultural Biotechnology, University of Florence.
[||] Hamburg Outstation.
[⊥] Russian Academy of Sciences.

(1) Parks, W. C.; Wilson, C. L.; Lopez-Boado, Y. S. *Nat. Rev. Immunol.* **2004**, *4*, 617–629.
(2) D'Alessio, S.; Fibbi, G.; Cinelli, M.; Guiducci, S.; Del Rosso, A.; Margheri, F.; Serrati, S.; Pucci, M.; Kahaleh, B.; Fan, P. S.; Annunziato, F.; Cosmi, L.; Liotta, F.; Matucci-Cerinic, M.; Del Rosso, M. *Arthritis Rheum.* **2004**, *50*, 3275–3285.
(3) Boire, A.; Covic, L.; Agarwal, A.; Jacques, S.; Sherifl, S.; Kuliopulos, A. *Cell* **2005**, *120*, 303–313.
(4) Page-McCaw, A.; Ewald, A. J.; Werb, Z. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 221–233.
(5) Aumailley, M.; Gayraud, B. *J. Mol. Med.* **1998**, *76*, 253–265.
(6) Knauper, V.; Will, H.; Lopez-Otin, C.; Smith, B.; Atkinson, S. J.; Stanton, H.; Hembry, R. M.; Murphy, G. *J. Biol. Chem.* **1996**, *271*, 17124–17131.
(7) Morrison, C. J.; Overall, C. M. *J. Biol. Chem.* **2006**, *281*, 26528–26539.
(8) Ra, H. J.; Parks, W. C. *Matrix Biol.* **2007**, *26*, 587–596.
(9) Overall, C. M.; Tam, E.; McQuibban, G. A.; Morrison, C.; Wallon, U. M.; Bigg, H. F.; King, A. E.; Roberts, C. R. *J. Biol. Chem.* **2000**, *275*, 39497–39506.
(10) Maskos, K.; Bode, W. *Mol. Biotechnol.* **2003**, *25*, 241–266.
(11) Andreini, C.; Banci, L.; Bertini, I.; Luchinat, C.; Rosato, A. *J. Proteome Res.* **2004**, *3*, 21–31.
(12) Van den Steen, P. E.; Grillet, B.; Opdenakker, G. *Glycobiol. Med.* **2005**, *564*, 45–55.
(13) Opdenakker, G.; Van den Steen, P. E.; Van Damme, J. *Trends Immunol.* **2001**, *22*, 571–579.
(14) Gronski, T. J.; Martin, R. L.; Kobayashi, D. K.; Walsh, B. C.; Holman, M. C.; Huber, M.; VanWart, H.; Shapiro, S. D. *J. Biol. Chem.* **1997**, *272*, 12189–12194.
(15) Mecham, R. P.; Broekelmann, T. J.; Fliszar, C. J.; Shapiro, S. D.; Welgus, H. G.; Senior, R. M. *J. Biol. Chem.* **1997**, *272*, 18071–18076.
(16) Murphy, G.; Allan, J. A.; Willenbrock, F.; Cockett, M. I.; O'Connell, J. P.; Docherty, A. J. *J. Biol. Chem.* **1992**, *267*, 9612–9618.

length active protein.[18,19] For this reason, it is often hypothesized that the HPX domain helps the local unwinding of the triple helix, in such a way that a single peptide strand can be accommodated in the active site of the CAT domain and cleaved.[20,21] It has been also hypothesized that a relative mobility of the HPX domain is necessary for this function.[22–25] This seems to be the case for MMP-9, for which small-angle X-ray scattering (SAXS) and atomic force microscopy (AFM) experiments indicate that the OG domain is able to lose its globular shape and transiently assume elongated structures, thereby allowing relative motion of the CAT and HPX domains.[26] On the other hand, X-ray structures of the full-length proenzyme forms of other MMPs lacking the OG domain display a well-defined structural relationship between the CAT and HPX domains, and this relationship is the same in the two different pro-MMPs studied (i.e., pro-MMP-1[25] (14 AA linker) and pro-MMP-2[27] (20 AA linker)).[25] The structures of active MMP-1,[28] and of its porcine ortholog,[29] are also compact and show a slightly different orientation of the HPX domain with respect to pro-MMP. This difference, although small, has been highlighted as evidence of the potential ability of the HPX domain to move with respect to the CAT domain.[25,28]

Here we have addressed the general problem of the relative conformational freedom of the two domains of MMPs in solution by NMR. NMR in solution is a powerful tool to investigate internal mobility of biomolecules[30–42] and can also provide precious information on interprotein and interdomain mobility.[43–49]

We selected MMP-12 (16 AA linker), for which an extended NMR assignment of the CAT domain is available, as well as high-resolution solid state and NMR structures. In this work, we have assigned the NMR signals of the HPX domain, solved its solution structure, and assigned the full-length protein. We have then obtained relaxation data (R_1, R_2, NOE) for the full-length protein and compared it with the same data for its isolated CAT and HPX domains. These data show that the two domains are not held rigidly to one another but must undergo independent motions. Residual dipolar couplings (RDC) on the full-length protein in the presence of an external orienting device were also obtained and found inconsistent with a rigid conformation of the protein.

Repeated attempts to crystallize full-length MMP-12 for X-ray diffraction finally yielded crystals of modest quality, which diffracted to 3 Å resolution. This low resolution was, however, largely sufficient to establish that the structure is less compact and the relative orientation of the two domains totally different, with respect to the four X-ray structures of MMP-1 and MMP-2 already described. Small-angle X-ray scattering (SAXS) data are consistent with the determined structure in the crystal representing the major conformation in solution but also point to a conformational mobility of the domains in MMP-12.[50]

Materials and Methods

Protein Expression. The cDNA encoding the G106-C470 sequence of full-length MMP-12 was generated by a polymerase chain reaction (PCR) from an IMAGE consortium clone using two synthetic oligonucleotides as primers. The cDNA obtained was cloned into the pET21a (Novagen) using the restriction enzymes *Nde I* and *Xho I* (New England BioLabs). One additional methionine at position 105 was present in the final expression product.

The expression vector encoding for the full-length protein (FL-MMP-12) was transformed into competent *Escherichia coli* BL21 (DE3) Codon Plus strain, and colonies were selected for Ampicillin and chloramphenicol resistance. Bacteria were grown in LB medium containing 34 μg/mL chloramphenicol and 50 μg/mL ampicillin in a shaker flask at 37 °C. Protein expression was induced with 0.5 mM IPTG at an $OD_{600} = 0.6$, and cell growth was continued for a further 5 h. For expression of ^{15}N and ^{13}C-enriched FL-MMP-12, the bacteria were grown in minimal medium containing ^{15}N enriched $(NH_4)_2SO_4$ and ^{13}C enriched glucose (Cambridge Isotope Laboratories). Cells were harvested by centrifugation and resuspended in a buffer containing 25% sucrose, 50 mM Tris-HCl (pH 8), 0.1 M NaCl, 0.2 M EDTA, 1 mM DTT. Five to ten milligrams of lysozyme were added to the resulting suspension and stirred for

(17) Overall, C. M.; McQuibban, G. A.; Clark-Lewis, I. *Biol. Chem.* **2002**, *383*, 1059–1066.
(18) Clark, I. E.; Cawston, T. E. *Biochem. J.* **1989**, *263*, 201–206.
(19) Ottl, J.; Gabriel, D.; Murphy, G.; Knauper, V.; Tominaga, Y.; Nagase, H.; Kroger, M.; Tschesche, H.; Bode, W.; Moroder, L. *Chem. Biol.* **2000**, *7*, 119–132.
(20) Chung, L. D.; Dinakarpandian, D.; Yoshida, N.; Lauer-Fields, J. L.; Fields, G. B.; Visse, R.; Nagase, H. *EMBO J.* **2004**, *23*, 3020–3030.
(21) Tam, E. M.; Moore, T. R.; Butler, G. S.; Overall, C. M. *J. Biol. Chem.* **2004**, *279*, 43336–43344.
(22) Gomis-Ruth, F. X.; Gohlke, U.; Betz, M.; Knauper, V.; Murphy, G.; Lopez-Otin, C.; Bode, W. *J. Mol. Biol.* **1996**, *264*, 556–566.
(23) de Souza, S. J.; Pereira, H. M.; Jacchieri, S.; Brentani, R. R. *FASEB J.* **1996**, *10*, 927–930.
(24) Overall, C. M. *Mol. Biotechnol.* **2002**, *22*, 51–86.
(25) Jozic, D.; Bourenkov, G.; Lim, N. H.; Visse, R.; Nagase, H.; Bode, W.; Maskos, K. *J. Biol. Chem.* **2005**, *280*, 9578–9585.
(26) Rosenblum, G.; Van den Steen, P. E.; Cohen, S. R.; Grossmann, J. G.; Frenkel, J.; Sertchook, R.; Slack, N.; Strange, R. W.; Opdenakker, G.; Sagi, I. *Structure* **2007**, *15*, 1227–1236.
(27) Morgunova, E.; Tuuttila, A.; Bergmann, U.; Isupov, M.; Lindqvist, Y.; Schneider, G.; Tryggvason, K. *Science* **1999**, *284*, 1667–1670.
(28) Iyer, S.; Visse, R.; Nagase, H.; Acharya, K. R. *J. Mol. Biol.* **2006**, *362*, 78–88.
(29) Li, J.; Brick, P.; Ohare, M. C.; Skarzynski, T.; Lloyd, L. F.; Curry, V. A.; Clark, I. M.; Bigg, H. F.; Hazleman, B. L.; Cawston, T. E.; Blow, D. M. *Structure* **1995**, *3*, 541–549.
(30) Powers, R.; Clore, G. M.; Stahl, S. J.; Wingfield, P. T.; Gronenborn, A. M. *Biochemistry* **1992**, *31*, 9150–9157.
(31) Szyperski, T.; Luginbuhl, P.; Otting, G.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **1993**, *3*, 151–164.
(32) Fischer, M. W. F.; Zeng, L.; Majumdar, A.; Zuiderweg, E. R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8016–8019.
(33) Ishima, R.; Torchia, D. A. *Nat. Struct. Biol.* **2000**, *7*, 740–743.
(34) Pfeiffer, S.; Fushman, D.; Cowburn, D. *J. Am. Chem. Soc.* **2001**, *123*, 3021–3026.
(35) Tolman, J. R.; Al-Hashimi, H. M.; Kay, L. E.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 1416–1424.
(36) Mulder, F. A. A.; Mittermaier, A.; Hon, B.; Dahlquist, F. W.; Kay, L. E. *Nat. Struct. Biol.* **2001**, *8*, 932–935.
(37) Eisenmesser, E. Z.; Bosco, D. A.; Akke, M.; Kern, D. *Science* **2002**, *295*, 1520–1523.
(38) Peti, W.; Meiler, J.; Brüschweiler, R.; Griesinger, C. *J. Am. Chem. Soc.* **2002**, *124*, 5822–5833.
(39) Bruschweiler, R. *Curr. Opin. Struct. Biol.* **2003**, *13*, 175–183.
(40) Lindorff-Larsen, K.; Best, R. B.; DePristo, M. A.; Dobson, C. M.; Vendruscolo, M. *Nature* **2005**, *433*, 128–132.
(41) Mittermaier, A.; Kay, L. E. *Science* **2006**, *312*, 224–228.
(42) Ferrage, F.; Pelupessy, P.; Cowburn, D.; Bodenhausen, G. *J. Am. Chem. Soc.* **2006**, *128*, 11072–11078.
(43) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. *Biochemistry* **1992**, *31*, 5269–5278.
(44) Fischer, M. W.; Losonczi, J. A.; Weaver, J. L.; Prestegard, J. H. *Biochemistry* **1999**, *38*, 9013–9022.
(45) Baber, J. L.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **2001**, *123*, 3953–3959.
(46) Jain, N. U.; Wyckoff, T. J. O.; Raetz, C. R. H.; Prestegard, J. H. *J. Mol. Biol.* **2004**, *343*, 1379–1389.
(47) Volkov, A. N.; Worrall, J. A. R.; Holtzmann, E.; Ubbink, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18945–18950.
(48) Ryabov, Y. E.; Fushman, D. *Magn. Reson. Chem.* **2006**, *44*, S143−S151.
(49) Ryabov, Y. E.; Fushman, D. *J. Am. Chem. Soc.* **2007**, *129*, 3315–3327.
(50) Bertini, I.; Calderone, V.; Cosenza, M.; Fragai, M.; Lee, Y.-M.; Luchinat, C.; Mangani, S.; Terni, B.; Turano, P. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5334–5339.

15−20 min at 4 °C. A buffer containing 2% Triton, 50 mM Tris-HCl (pH 8), 0.1 M NaCl, 0.2 M EDTA, and 1 mM DTT was added, and the suspension was sonicated (7−8 30 s cycles) and centrifuged at 40 000 rpm for 20 min at 4 °C. The pellets were resuspended in 6 M Urea, 20 mM Tris-HCl (pH 8) and centrifuged. The resulting inclusion bodies were solubilized in 6 M Gdn-HCl, 20 mM Tris-HCl (pH 8), 10 mM DTT, and 20 mM cystamine and stirred overnight at 4 °C. The insoluble material was removed by centrifuging at 9000 rpm for 10 min. Protein molecular weight and purity was checked on 15% gel by SDS-PAGE. The F171D, E219A mutant of FL-MMP-12 was produced using the quick-change site-directed mutagenesis kit (Qiagen), and the expression and purification of the protein and of its ^{15}N- and ^{13}C−^{15}N-enriched versions completed using the same procedure described above.

Refolding was carried out by a serial dilution method.[51] The protein was diluted to a final concentration of 0.13 mg/mL in 6 M Gdn-HCl, 20 mM Tris-HCl (pH 8), 1 mM DTT, and 0.05% Brij-35, stirred at 4 °C for 30−60 min, and dialyzed overnight against a buffer containing 20 mM Tris (pH 7.2), 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.15 M NaCl, 5 mM β-mercaptoethanol, 1 mM 2-hydroxyethyl disulfide, and 0.05% Brij-35. The refolded protein was then purified using a Sepharose column (Amersham) with a buffer containing 20 mM Tris (pH 7.2), 10 mM $CaCl_2$, 0.3 M NaCl and 0.1 M acetohydroxamic acid (AHA). The protein solution was concentrated up to 5 mL and purified by size exclusion chromatography using a High Load 16/60 Superdex 75 (Amersham Biosciences) and eluted with 20 mM Tris pH 7.2, 10 mM $CaCl_2$, 0.3 M NaCl, and 0.2 M AHA. The eluted fractions were checked for purity on 15% gel by SDS-PAGE, and those containing the FL-MMP-12 protein were pooled and concentrated.

Samples of cadmium(II) substituted FL-MMP-12 protein were prepared by exhaustive dialysis against a buffer containing 20 mM Tris pH 7.2, 10 mM $CaCl_2$, 0.3 M NaCl, 0.2 M AHA, and 0.3 mM of $CdCl_2$.[52] Equimolar concentrations of *N*-isobutyl-*N*-[4methoxyphenylsulfonyl] glycyl hydroxamic acid (NNGH) were added to the samples to further increase the protein stability.

The cDNA encoding for the HPX domain (E278-C470) was generated by polymerase chain reaction and cloned into pET21a, using *Nde I* and *Xho I* as restriction enzymes. The expression vector was then transformed into competent *E. coli* BL21 (DE3) Gold strain, and the colonies were selected for Ampicillin resistance. Protein refolding for both nonlabeled and for ^{13}C and/or ^{15}N enriched samples was carried out by using the same protocols previously described for the preparation of FL-MMP-12 samples. Samples of the zinc(II) catalytic domain (F171D mutant) were prepared as previously described.[50]

NMR Measurements and Solution Structure Calculations. The experiments for the structure calculation and mobility measurements of the isolated HPX domain were performed on protein samples at concentrations ranging between 0.5 and 0.7 mM (pH 7.2). For FL-MMP-12, all NMR experiments were performed on samples at a concentration of 0.5 mM (pH 7.2).

All NMR experiments were performed at 298 K and acquired on Bruker AVANCE 900, AVANCE 800, AVANCE 700 and DRX 500 spectrometers. All instruments but one are equipped with triple resonance CRYO-probes. The 700 MHz spectrometer is equipped with a triple resonance (TXI) 5 mm probe with a *z*-axis pulse field gradient.

All spectra were processed with the Bruker TOPSPIN software packages and analyzed by the program CARA (Computer Aided Resonance Assignment, ETH Zürich).[53]

The backbone resonance assignment was obtained by the analysis of HNCA, HNCO, HN(CA)CO, HNCACB, and CBCA(CO)NH spectra performed at 900 MHz. The assignment of the aliphatic side-chain resonances was performed through the analysis of 3D (H)CCH-TOCSY spectra at 500 MHz, together with 3D ^{15}N- and ^{13}C-NOESY-HSQC spectra at 900 MHz. The obtained assignments are reported in Tables S1 and S2 for the full length protein and tables S3 and S4 for the hemopexin domain. $^3J_{HNH\alpha}$ coupling constants were determined through the HNHA experiment at 500 MHz. Backbone dihedral φ angles were independently derived from $^3J_{HNH\alpha}$ coupling constants through the appropriate Karplus equation.[54] Backbone dihedral ψ angles for residue *i*-1 were also determined from the ratio of the intensities of the $d_{\alpha N}$(i-1,i) and $d_{N\alpha}$(i,i) NOEs present on the ^{15}N(*i*) plane of residue *i* obtained from the ^{15}N-edited NOESY-HSQC spectrum.

3D ^{15}N- and ^{13}C-enriched NOESY-HSQC cross peak intensities were integrated using the integration routine implemented in CARA and converted into interatomic upper distance limits by the program CALIBA.[55]

The protein assignment and the mobility measurements on FL-MMP-12 where performed on the NNGH-inhibited, cadmium(II)-substituted Phe171Asp/Glu219Ala mutant, due to its high stability to the self-hydrolysis. Mobility measurements on the catalytic domain where performed on the NNGH-inhibited, zinc(II) form of the Phe171Asp mutant.[50]

Residual dipolar couplings have been measured on FL-MMP-12 in the presence of an external orienting medium constituted by a binary mixture of $C_{12}E_5$ (penta-ethyleneglycol dodecyl ether, Fluka) and neat *n*-hexanol (Fluka), with a molar ratio $C_{12}E_5$/*n*-hexanol of 0.96 and with a $C_{12}E_5$/water ratio of 5 wt %.[56] One-bond ^{1}H−^{15}N coupling constants were measured at 298 K and 900 MHz by using the IPAP method.[57,58] Two-hundred fifty-nine rdc values could be measured that ranged from −46 to +25 Hz. Of them, only those rdc values corresponding to residues experiencing neither mobility nor large rmsd (140 residues, mostly in α or β secondary structures, see Table S5, Supporting Information) have been used for structure calculations and to investigate the reciprocal mobility of the two domains.

The program DYANA[59] was used to calculate a family of 1600 structures of the isolated HPX domain starting from randomly generated conformers in 20 000 annealing steps. The solution structure statistics are reported in Table S6 (Supporting Information). The family was energy-minimized by iterative cycles of DYANA with the program FANTAORIENT.[60] The quality of the structures calculated by DYANA can be assessed by a properly defined energy function (target function) proportional to the squared deviations of the calculated constraints from the experimental ones, plus the standard covalent and nonbonded energy terms. Structure calculation statistics and the structural quality were evaluated using the program PROCHECK_NMR.[61]

R_1, R_2, and NOE Measurements. The experiments for the determination of ^{15}N longitudinal and transverse relaxation rates and ^{1}H−^{15}N NOE were recorded at 298 K and 700 MHz on ^{15}N-enriched samples. The ^{15}N longitudinal relaxation rates (R_1) were measured using a sequence modified to remove cross correlation effects during the relaxation delay.[62] Inversion−recovery times

(51) Bertini, I.; Calderone, V.; Fragai, M.; Luchinat, C.; Mangani, S.; Terni, B. *Angew. Chem. Int. Ed.* **2003**, *42*, 2673–2676.
(52) Bertini, I.; Fragai, M.; Lee, Y.-M.; Luchinat, C.; Terni, B. *Angew. Chem. Int. Ed.* **2004**, *43*, 2254–2256.
(53) Keller, R. L. J. The Computer Aided Resonance Assignment Tutorial, 1.3; Cantina: Verlag, 2004.
(54) Vuister, G. W.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 7772–7777.
(55) Guntert, P.; Braun, W.; Wüthrich, K. *J. Mol. Biol.* **1991**, *217*, 517–530.
(56) Rückert, M.; Otting, G. *J. Am. Chem. Soc.* **2000**, *122*, 7793–7797.
(57) Otting, G.; Ruckert, M.; Levitt, M. H.; Moshref, A. *J. Biomol. NMR* **2000**, *16*, 343–346.
(58) Ottiger, M.; Delaglio, F.; Bax, A. *J. Magn. Reson.* **1998**, *131*, 373–378.
(59) Guntert, P.; Mumenthaler, C.; Wüthrich, K. *J. Mol. Biol.* **1997**, *273*, 283–298.
(60) Bertini, I.; Luchinat, C.; Parigi, G. *Concepts Magn. Reson.* **2002**, *14*, 259–286.
(61) Laskowski, R. A.; Rullmann, J. A. C.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. *J. Biomol. NMR* **1996**, *8*, 477–486.
(62) Kay, L. E.; Nicholson, L. K.; Delaglio, F.; Bax, A.; Torchia, D. A. *J. Magn. Reson.* **1992**, *97*, 359–375.

ranging between 2.5 and 3000 ms, with a recycle delay of 3.5 s, were used for the experiments. The ^{15}N transverse relaxation rates (R_2) were measured using a CPMG sequence[62,63] with delays ranging between 8.5 and 237.4 ms for the CAT domain, between 8.5 and 203.5 ms for the HPX domain, and finally between 8.5 and 135.7 ms for the FL-MMP-12 protein with a refocusing delay of 450 μs. The relaxation data are reported in Table S7 (Supporting Information). R_1 and R_2 data measured on the full length protein were found nosier and less uniform with respect to those of the single catalytic and hemopexin domains. This is related to the overlap of the signals in a so large protein and to the relative low solubility of the full length construct.

Crystallization, Data Collection, and X-ray Structure Determination. Crystals of the zinc(II) Phe171Asp/Glu219Ala FL-MMP-12 were obtained using the vapor diffusion technique at 20 °C from a solution containing 0.1 M Tris-HCl, 30% PEG-8000, 200 mM AHA, pH 8.0. The final protein concentration was 0.7 mM.

The data collection was carried out in-house, using a PX-Ultra copper sealed tube source (Oxford Diffraction) equipped with an Onyx CCD detector. The data set was collected at 100 K and the crystals used for data collection were cryo-cooled using a solution containing 10% ethylene glycol in the mother liquor.

The crystal diffracted to 3.0 Å resolution; it belongs to space-group C2 with one molecule in the asymmetric unit, a solvent content of about 50%, and a mosaicity of about 1.0°.

The data was processed using the program MOSFLM[64] and scaled using the program SCALA[65] with the TAILS and SECONDARY corrections on (the latter restrained with a TIE SURFACE command) to achieve an empirical absorption correction. The structure was solved using the molecular replacement technique in two following steps; in the first step, the model used to find the correct orientation of the catalytic part of the structure was 1Y93 whereas the one used for the hemopexin domain was the same domain in proMMP-1 structure (1SU3). In both steps, inhibitors, water molecules, and ions were omitted from the models.

The correct orientation and translation of the molecules within the crystallographic unit cell was determined with standard Patterson search techniques[66,67] as implemented in the program MOLREP.[68,69] The isotropic refinement was carried out using REFMAC5[70] and default weights for the crystallographic term and the geometrical term were used.

In between the refinement cycles, the models were subjected to manual rebuilding by using XtalView.[71] The same program was used to model AHA molecule. Water molecules have been added by using the standard procedures within the ARP/WARP suite.[72] The stereochemical quality of the refined models was assessed using the program PROCHECK.[73] The Ramachandran plot is of average quality for such a resolution structure. Of the seven violations four (Asp303, Ser311, Lys315, and Asn363) are located in the hemopexin domain; it is interesting to compare these outliers to the equivalent residues in MMP-1 (1SU3) which has a higher resolution (2.2 Å). Asp299 (corresponding to Asp303 in MMP-12) is an outlier as well, and in the place of Ser311 there is a proline residue (Pro307) which has of course a unique stereochemistry; furthermore in the place of Asn363 there is Gly359 in MMP-1, which again is unique from the stereochemical point of view. Therefore, three out of the four residues that are in disallowed regions in the HPX domain may actually assume peculiar positions in the 3-D structure. Table S8 (Supporting Information) reports the data collection and refinement statistics.

SAXS Experiments and Data Analysis. Synchrotron X-ray scattering data from solutions of the NNGH-inhibited, cadmium(II)-substituted Phe171Asp/Glu219Ala double mutant of FL-MMP-12 were collected on the X33 beamline of the EMBL (DESY, Hamburg),[74] using a MAR345 image plate detector. The scattering patterns were measured with a 2 min exposure time for several solute concentrations in the range from 0.8 to 8.3 mg/mL. To check for radiation damage, two 2 min exposures were compared, and no changes were detected. Using the sample−detector distance of 2.7 m, a range of momentum transfer of $0.01 < s < 0.5$ Å^{-1} was covered ($s = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle, and $\lambda = 1.5$ Å is the X-ray wavelength). The data were processed using standard procedures and extrapolated to infinite dilution using the program PRIMUS.[75] The forward scattering, *I(0)*, and the radius of gyration, R_g, were evaluated using the Guinier approximation,[76] assuming that at very small angles ($s < 1.3/R_g$) the intensity is represented as $I(s) = I(0) \exp(-(s^2R_g^2)/3$. The values of $I(0)$ and R_g, as well as the maximum dimension, D_{max}, and the interatomic distance distribution functions, ($p(r)$), were also computed using the program GNOM.[77] The molecular mass of FL-MMP-12 was evaluated by comparison of the forward scattering with that for a reference solution of bovine serum albumin (66 kDa).

The scattering from the high resolution models was computed using the program CRYSOL.[78] Given the atomic coordinates, the program either predicts the theoretical scattering pattern or fits the experimental intensity by adjusting the excluded volume of the particle and the contrast of the hydration layer to minimize the discrepancy

$$x^2 = \frac{1}{N-1}\sum_i \left[\frac{I_{exp}(S_j) - cI_{calc}(S_j)}{\sigma(S_j)}\right]^2 \quad (1)$$

where N is the number of experimental points, c is a scaling factor, $I_{exp}(s_j)$ and $I_{calc}(s_j)$ are the experimental and calculated intensities, respectively, and $\sigma(s_j)$ is the experimental error at the momentum transfer s_j.

To assess the conformational variability of MMP-12, an ensemble optimization method (EOM) was used,[79] allowing for coexistence of multiple conformations in solution. About 5000 randomized models of MMP-12 differing by the conformation of the interdomain linker were generated using the program DYANA starting from randomly generated conformers of the full-length protein where only the dihedral angles of the linker region were left free to vary. These models formed a pool of possible structures, for which the scattering patterns were computed by CRYSOL. The EOM program employs a genetic algorithm to select from the pool a small number (usually about 20) of representative structures such that the average scattering from the selected ensemble fits the experimental data. Multiple runs of EOM were performed and the results were averaged to provide quantitative information about

(63) Peng, J. W.; Wagner, G. *Methods Enzymol.* **1994**, *239*, 563–596.
(64) Leslie, A. G. W. In *Molecular data processing*; Moras, D., Podjarny, A. D., Thierry, J.-C., Eds.; Oxford University Press: Oxford, 1991; pp 50−61.
(65) Evans, P. R. *Joint CCP4 and ESF-EACBM Newsletter* **1997**, *33*, 22–24.
(66) Rossmann, M. G.; Blow, D. M. *Acta Crystallogr.* **1962**, *15*, 24.
(67) Crowther, R. A. *The Molecular Replacement Method*; Rossmann, M. G., Ed.; Gordon & Breach: New York, 1972.
(68) Vagin, A.; Teplyakov, A. *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.
(69) Vagin, A.; Teplyakov, A. *Acta Crystallogr. D: Biol. Crystallogr.* **2000**, *56*, 1622–1624.
(70) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr.* **1997**, *D53*, 240–255.
(71) McRee, D. E. *J. Mol. Graphics* **1992**, *10*, 44–47.
(72) Lamzin, V. S. *Acta Crystallogr. D: Biol. Crystallogr.* **1993**, *49*, 129–147.
(73) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
(74) Roessle, M. W.; Klaering, R.; Ristau, U.; Robrahn, B.; Jahn, D.; Gehrmann, T.; Konarev, P. V.; Round, A.; Fiedler, S.; Hermes, S.; Svergun, D. I. *J. Appl. Crystallogr.* **2007**, *40*, s190–s194.
(75) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. *J. Appl. Crystallogr.* **2003**, *36*, 1277–1282.
(76) Guinier, A. *Ann. Phys.* **1939**, *12*, 161–237.
(77) Svergun, D. I. *J. Appl. Crystallogr.* **1992**, *25*, 495–503.
(78) Svergun, D. I.; Barberato, C.; Koch, M. H. J. *J. Appl. Crystallogr.* **1995**, *28*, 768–773.
(79) Bernado, P.; Mylonas, E.; Petoukhov, M. V.; Blackledge, M.; Svergun, D. I. *J. Am. Chem. Soc.* **2007**, *129*, 5656–5664.

the flexibility of the protein in solution (in particular, about the R_g distribution in the selected ensembles).

Results and Discussion

Design and Production of a Stable Full-Length MMP-12. Several MMPs, both as isolated catalytic domains (CAT) or full-length (FL) proteins, are relatively unstable *in vitro* due to self-proteolysis (while the inactive isolated hemopexin domain (HPX) is stable). In the case of MMP-12, the CAT domain can be stabilized by a Phe171Asp mutation that prevents self-proteolysis and increases the solubility. This mutation is far both from the active site and from the CAT-HPX interface (see later) and does not perturb the catalytic activity of the enzyme. Therefore, a FL construct with the Phe171Asp mutation was initially produced. The resulting protein was well folded and fully active but still sensitive to the self-proteolysis at the linker region even in the presence of the inhibitor NNGH.

To increase the stability, the catalytically relevant Glu219 was mutated to an Ala. This mutation had been shown to decrease the catalytic activity of the isolated CAT domain of MMPs by two-3 orders of magnitude,[80] whereas the three-dimensional structure of the domain is fully retained.[81] The Phe171Asp/Glu219Ala double mutant of the FL protein was stable for several days.

Finally, we substituted the catalytic zinc(II) ion, responsible for the residual activity, with a cadmium(II) ion. The resulting derivative, complexed with NNGH, resulted stable to proteolysis, showing no trace of cleaved CAT and HPX domains in the gel after three weeks of NMR measurements. The analysis of the HSQC spectra revealed that no structural alteration were caused by the replacement of zinc(II) with the cadmium(II) ion.

The NNGH-inhibited, cadmium(II)-substituted Phe171Asp/Glu219Ala double mutant of FL-MMP-12 was therefore used in all NMR experiments reported hereafter.

Low-Resolution X-Ray Structure of FL-MMP-12 and Comparison with Existing Structural Information on FL-MMP. Crystals diffracting at 3.0 Å were obtained for the AHA-inhibited, zinc(II) Phe171Asp/Glu219Ala double mutant of FL-MMP-12.

The X-ray structure of this construct was solved from diffraction data obtained in-house and deposited in PDB (PDB code: 3BA0). Despite the low resolution, the electron density was of generally good quality throughout the entire molecule, the only exception being residues 271−274 in the middle of the linker region (see below), which spans from Asp264 to Pro279. The overall structure of the CAT domain is very similar to the high resolution structure of the isolated CAT domain.[50] The mutations are both clearly visible; the active site contains both the native catalytic zinc ion and the structural zinc ion. Three calcium atoms are also present in the structure of the CAT domain. Finally, a AHA molecule bound to the zinc ion in the usual geometry[50] is present at the active site. The HPX domain has the expected four-blade propeller, hemopexin-like fold previously observed in other HPX domains of MMPs.[22,25,27,29,82–85] The disulfide bridge between Cys282 and Cys470 is clearly present, and a calcium ion is bound in the central region of the domain.

The relative orientation of the two domains could be unambiguously determined from the 3 Å X-ray data (Figure 1A). The two domains are in contact through a relatively small surface (∼165 Å^2) which includes a possible salt bridge between His112 of CAT domain and the C-terminal of Cys470 in the HPX domain, and van der Waals interactions between residue 113 of CAT and residues 284 and 463 on the HPX side. The first part of the linker (residues 264, 265, 266, 270, 271, and 273) is in touch with the CAT domain (residues 112, 143, 144, 249, 259, and 263) through a surface of 902 Å^2 and the last part (residues 277 and 279) with the HPX domain (residues 280, 307, and 470) through a surface of 571 Å^2 (Figure 1A). Furthermore, the linker residue Lys266, besides being in contact with the CAT domain, forms a salt bridge with the C-terminal of HPX domain. The central part of the linker (residues 271−274) shows a rather poor electron density, but its structure can be reasonably modeled to match the arrangement of the initial and final linker residues. It is apparent that the whole molecule is held together mainly by interactions between each single domain and the linker rather than between the two domains. The presence of the intact linker was confirmed by a gel experiment on a protein solution obtained after redissolving crystals collected from the same wells of those used for X-ray.

Figure 1C shows a superposition of the X-ray structures of human proMMP-1 and proMMP-2, the only FL-proMMPs structures available to date, where the prodomain and the fibronectin domains (in case of MMP-2) have been omitted for clarity. The CAT-HPX interface is very similar in the two proteins and relatively extended (ca. 753 and 866 Å^2, respectively, as opposed to ca. 165 Å^2 in FL-MMP-12), suggestive of a stable domain-domain interaction. Figure 1D shows a superposition of the structures of active human FL-MMP-1[28] and of its porcine ortholog.[29] Again, the two structures are very similar, and the CAT-HPX interface is again extended (ca. 735 and 747 Å^2, respectively). The relative orientation of the HPX domain is slightly different in these active forms with respect to the pro-enzyme forms (cfr. Figure 1D and 1C, where the CAT domain has the same orientation), as previously noticed.[28]

Relative to the CAT domain, the HPX domain in FL-MMP-12 lies at about 120° with respect to its orientation in the structures of MMP-1 and MMP-2 (Figure 1E), and the two domains have a less compact arrangement (cfr. Figure 1A and 1B). The interface region is completely different, both viewed from the CAT and from the HPX domain, is sensibly smaller (ca. 165 vs 735−866 Å^2), and the complementarity between the surfaces of the two domains is poor with respect to that observed in the experimental structures of FL-MMP-1 and FL-MMP-2.

NMR Characterization of FL-MMP-12. The NNGH-inhibited, cadmium(II)-substituted Phe171Asp/Glu219Ala double mutant of FL-MMP-12 (FL-MMP-12 hereafter) yields ^{15}N−^{1}H HSQC spectra of surprisingly good quality for a protein of 42 000 Da (see Figure 2A) (spectra of this type were obtained on all constructs of FL-MMP-12 described before).

Figure 2B shows the ^{15}N−^{1}H HSQC spectrum of FL-MMP-12 (in black) superimposed to those of the isolated CAT domain (in green), prepared as previously reported, and of the isolated HPX domain (in red, see below). The similarity is striking: except for a number of peaks that are not present in either the

(80) Cha, J.; Auld, D. S. *Biochemistry* **1997**, *36*, 16019–16024.
(81) Bertini, I.; Calderone, V.; Fragai, M.; Luchinat, C.; Maletta, M.; Yeo, K. J. *Angew. Chem., Int. Ed.* **2006**, *45*, 7952–7955.
(82) Morgunova, E.; Tuuttila, A.; Bergmann, U.; Tryggvason, K. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7414–7419.
(83) Libson, A. M.; Gittis, A. G.; Collier, I. E.; Marmer, B. L.; Goldberg, G. I.; Lattman, E. E. *Nat. Struct. Biol.* **1995**, *2*, 938–942.
(84) Gohlke, U.; Gomis-Ruth, F. X.; Crabbe, T.; Murphy, G.; Docherty, A. J.; Bode, W. *FEBS Lett.* **1996**, *378*, 126–130.
(85) Cha, H.; Kopetzki, E.; Huber, R.; Lanzendorfer, M.; Brandstetter, H. *J. Mol. Biol.* **2002**, *320*, 1065–1079.

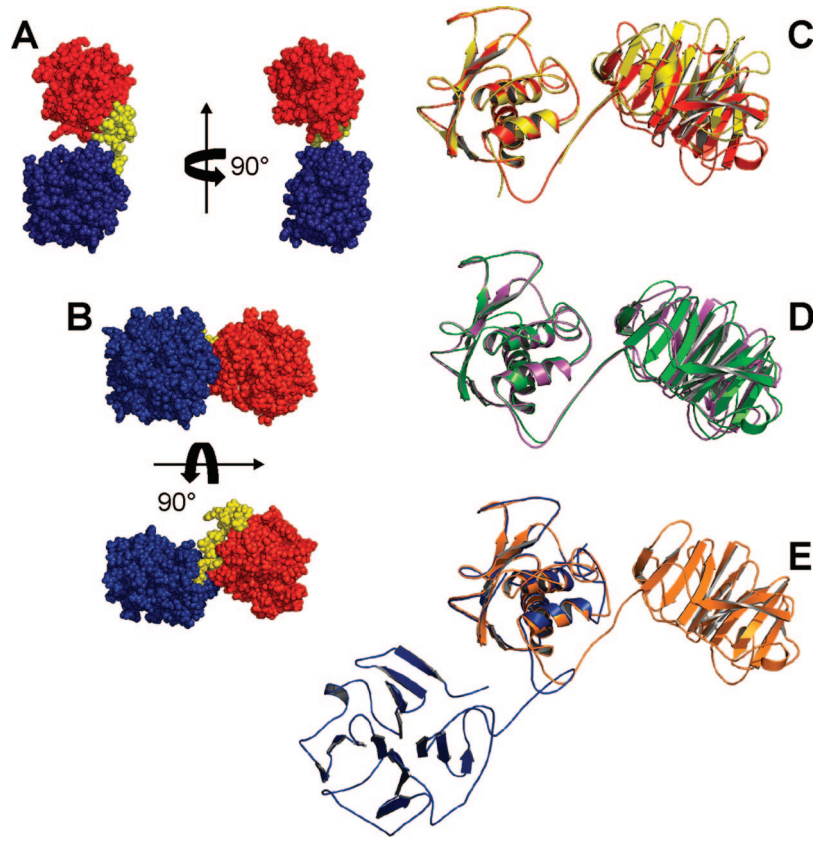


Figure 1. Space-filling representations of the X-ray structures of FL-MMP-12 (A, present work) and FL-MMP-1 (B).[28] Superimposition of the X-ray structures of (C) human pro-MMP-1[25] (red) and pro-MMP-2[27] (yellow) (where the prodomain of both and the fibronectin domains of MMP-2 have been omitted for clarity), (D) human[28] (violet) and porcine[29] (green) FL-MMP-1, and (E) human FL-MMP-1[28] (violet) with the present FL-MMP-12 (blue).

isolated CAT and HPX domains, which must therefore belong to the linker, all other FL-MMP-12 peaks are superimposable, or nearly superimposable, to a peak of either the isolated CAT or HPX domain. Furthermore, the peaks of FL-MMP-12 were only marginally broader that the corresponding peaks in the isolated domains.

The ^{13}C, ^{15}N, and ^{1}H spectral assignment of the CAT domain was available,[50] whereas that of the HPX domain was performed during this research. Also the solution structure of the HPX domain, up to now unavailable for MMP-12, was experimentally solved here (see Table S6 and Figure S1, Supporting Information) and deposited in PDB (PDB code: 2JXY). Ribbon representations of the solution and crystal structures are reported in Figure S1. The relatively high rmsd between the two structures (BB = 1.88 Å, secondary structure elements = 1.16 Å) is within the indetermination of both the X-ray (3.0 Å resolution) and NMR structures (BB rmsd = 1.38 Å, heavy atom rmsd = 2.13 Å).

The $^{15}N-^{1}H$ HSQC spectrum of FL-MMP-12 could be largely assigned by direct comparison with the isolated CAT and HPX spectra. Nevertheless, 97% of the nonlinker residues were reassigned by recording the same set of 3D spectra used for the structure of the isolated domains. Despite the crowding, the spectra were of good enough quality to assign most of the backbone atoms of the two domains as well as a large fraction of side-chain atoms (Table S1, Supporting Information). 3D NOESY spectra were also recorded. The NOESY patterns were largely the same as those observed in the isolated CAT and HPX domains, indicating that the domain structures are indeed unaltered in the FL protein, as it was already apparent from the comparison of the HSQC spectra. No interdomain NOEs could be found. The largest difficulties in the assignment were encountered in the linker region. Although several of the linker $^{15}N-^{1}H$ HSQC peaks could be identified as the extra peaks present in the FL protein spectra and not corresponding to peaks of any of the two domains, several of the necessary sequential connectivities were missing in the 3D spectra. In addition, very few NOEs could be seen for these peaks. A protonless CON experiment[86] allowed us to identify two of the linker prolines. Ten linker peaks could be identified, and four of them sequence-specifically assigned (Table 1).

Relaxation Data. The 700 MHz R_1 and R_2 data for FL-MMP-12 as well as for the isolated CAT and HPX domains are shown in Figure 3A−D. R_1 and R_2 values can be accurately estimated from the atomic coordinates of a macromolecule of known structure assuming a rigid-body hydrodynamics by using computer programs like HYDRONMR.[87] In this study, the X-ray structure of the CAT domain (PDB code: 1Y93),[50] the present solution structure of the HPX domain, and the present X-ray structure of the full length protein (Figure 1E) were used as input files in HYDRONMR. As shown in Figure 3, while for the isolated domains the experimental data nicely match the calculated values, for FL-MMP-12 sizably faster longitudinal and sizably slower transverse relaxation rates with respect to the calculated values were measured. These differences are clearly statistically significant, despite the larger errors due to the signal overlap and lower signal-to-noise ratios in the spectra

(86) Bermel, W.; Bertini, I.; Duma, L.; Emsley, L.; Felli, I. C.; Pierattelli, R.; Vasos, P. R. *Angew. Chem., Int. Ed.* **2005**, *44*, 3089–3092.
(87) de la Torre, J. G.; Huertas, M. L.; Carrasco, B. *J. Magn. Reson.* **2000**, *147*, 138–146.

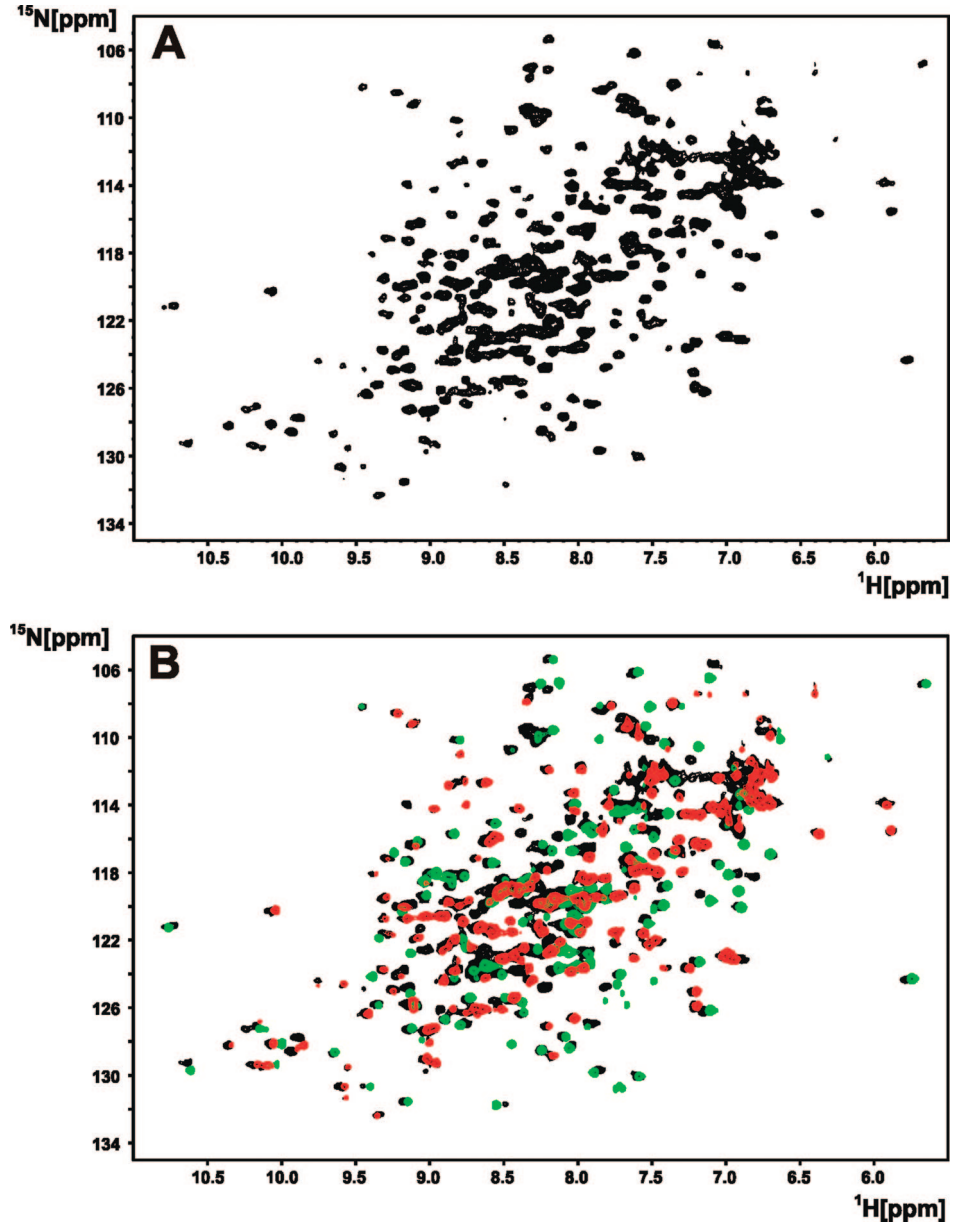


Figure 2. 900 MHz $^{15}N-^{1}H$ HSQC spectra of the uniformly ^{15}N-labelled NNGH-inhibited, cadmium(II)-substituted Phe171Asp/Glu219Ala double mutant of FL-MMP-12 (A), and superimposed spectra of FL-MMP-12 (black), CAT-MMP-12 (green), and HPX-MMP-12 (red) (B).

Table 1. ^{1}H, ^{15}N, and ^{13}C Chemical Shifts for the Identified Linker Residues of FL-MMP-12

		HN	N	CA	HA	CB	other	NOE
264	**D**	7.88	116.40	48.44	4.55	36.50	HB2 3.77, HB3 2.53	
272	**P**			60.17		29.20	CG 24.29, CD 46.96	
273	**N**	8.38	119.79	48.44	4.78	35.86	HB2 2.72, HB3 2.56	0.25
279	**P**			59.92	4.20	29.18	HB2 2.12, HB3 1.67, HG2 1.87, HG3 1.5, HD2 3.18, HD3 3.84	
L1		8.28	120.71	53.34	4.17	29.97	QB 1.78, CG 38.97, QG 2.87	
L2		8.17	109.42	57.03	4.83	39.62	HB2 3.05, HB3 2.60	
L3		7.89	122.58	49.62	4.18			
L4		8.16	121.91	53.24	4.19	30.91		0.28
L5		7.34	107.95	56.64				0.35
L6		8.37	123.69	51.65				

of FL-MMP-12. The R_1 and R_2 values measured on FL-MMP-12 are intermediate between the expectation from the isolated domains and a rigid full length structure. At the same time, the NOE values for FL-MMP-12 (Figure 3E) demonstrate that the single domains (catalytic and hemopexin) forming the full length protein behave as rigid bodies. Therefore, relaxation data can be collectively taken as evidence that the full length protein does not exhibit a rigid body hydrodynamics but experience flexibility. Such flexibility must depend on the presence of a flexible linker that permits sizable reciprocal mobility of the two domains on a time scale that is faster than the reorientation time of the whole molecule.

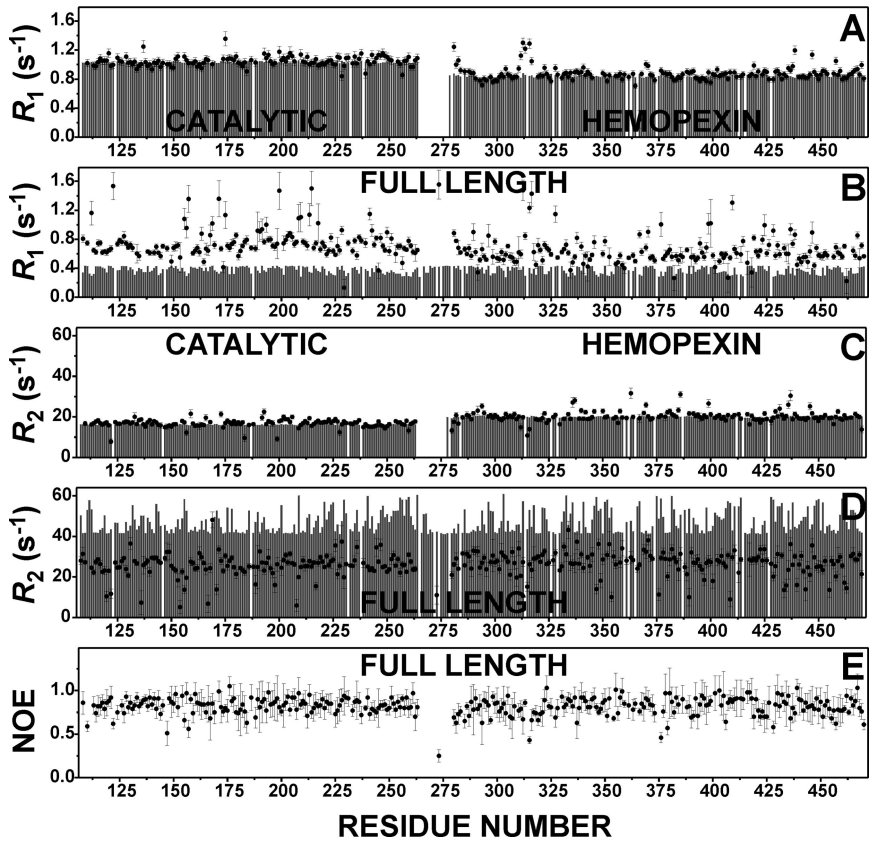


Figure 3. Calculated (grey bars) and experimental (filled circles) backbone ^{15}NH R_1 (A, B) and R_2 (C, D) values for the isolated CAT and HPX domains (A, C) and for the full-length protein (B, D). Although the agreement between experimental and calculated R_1 and R_2 values for the isolated domains is excellent, for the full-length protein the experimental R_1 values are sizably larger (B) and the R_2 values sizably smaller (D) than the ones calculated for the rigid X-ray structure. Experimental NOE values for the full-length protein (E).

Indeed, three linker peaks (N273, L4, L5) display grossly altered relaxation values, and particularly ^{15}N−^{1}H NOEs that are sizably smaller than expected (0.25, 0.28, and 0.35 respectively; Table 1). This is another indication that the linker is at least partially involved in some fast conformational rearrangement, consistent with a sizable degree of reciprocal mobility of the two domains. Incidentally, similar NOE values were observed in the linker region of calmodulin, a two-domain protein known to sample an extremely large conformational space,[43,45,50,88] and more recently in the two-domain xylanase Cex.[89] All R_1 and R_2 data on FL-MMP-12 and on its isolated CAT and HPX domains, together with NOE data, are reported in the Supporting Information (Table S7).

Residual Dipolar Couplings. Residual dipolar couplings (RDC) in the presence of the external orienting device C12E5/hexanol[56] have been measured. They can be fitted very well to the structures of the two isolated domains (Figure 4A,B) separately, but with sensibly different orientation tensor values. The data are, instead, in striking disagreement with the solid state X-ray structure (Figure 4C). Figure 4D shows that the agreement is modest also for any of the four rigid two-domain structures that can be obtained by fitting both domains to a single orientation tensor. None of these four structures bears a resemblance with the X-ray structure in terms of relative domain orientation (Figure S2, Supporting Information). These structures are also different from any other X-ray structure of FL-MMPs. On this basis, and on the basis of the relaxation data, these solutions are discarded.

SAXS Experiments. The processed X-ray scattering pattern from FL-MMP-12 presented in Figure 5 yields a molar mass estimate of 40 ± 4 kDa, compatible with that calculated from the sequence (42.5 kDa), indicating that the protein is monomeric in solution. The experimental radius of gyration R_g and maximum size D_{max} are 31 ± 1 Å and 110 ± 10 Å, respectively. These values significantly exceed the parameters calculated from the X-ray structure of FL-MMP-1 (R_g = 25 Å, D_{max} = 85 Å), while they are in better agreement with those computed from the less compact X-ray structure of FL-MMP-12 determined in the present study (R_g = 29 Å, D_{max} = 95 Å). Moreover, the scattering pattern calculated from the FL-MMP-1 model using CRYSOL[78] fails to fit the experimental data (discrepancy χ = 5.8, curve 2 in Figure 5).

The scattering curve computed from the present FL-MMP-12 structure displays a much better agreement to the experiment (χ = 2.5, curve 3 in Figure 5), but still displays some systematic deviations. For flexible MMP-12, neither individual models, nor averaging over the random pool allowed one to fit the SAXS data. The representative ensembles selected to fit the data give information about the preferable conformations of the protein. Given the potential conformational flexibility of MMP-12 in solution suggested by the NMR data, an alternative analysis approach was applied, allowing for the coexistence of multiple protein configurations. A large number of models was generated

(88) Bertini, I.; Del Bianco, C.; Gelis, I.; Katsaros, N.; Luchinat, C.; Parigi, G.; Peana, M.; Provenzani, A.; Zoroddu, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 6841–6846.
(89) Poon, D. K. Y.; Withers, S. G.; McIntosh, L. P. *J. Biol. Chem.* **2007**, *282*, 2091–2100.

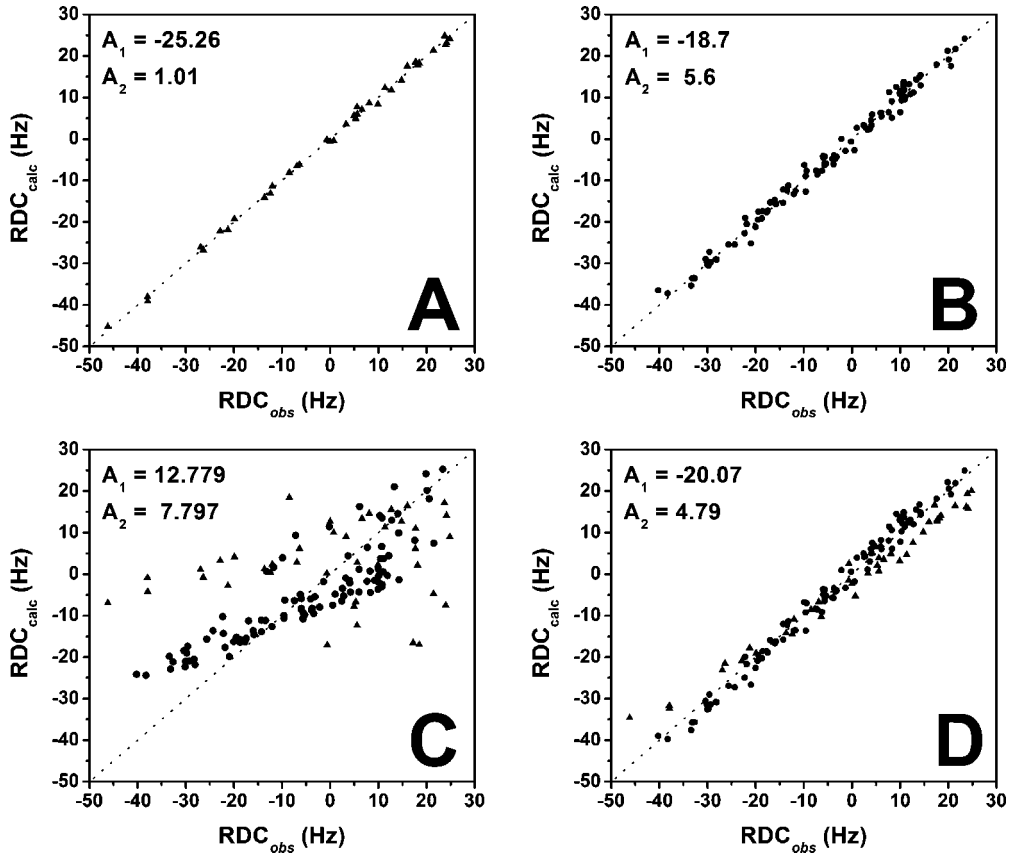


Figure 4. Best fit vs experimental 900 MHz RDC values for the NNGH-inhibited, cadmium(II)-substituted Phe171Asp/Glu219Ala double mutant of FL-MMP-12 in a binary mixture of $C_{12}E_5$ (penta-ethyleneglycol dodecyl ether) and neat *n*-hexanol. The separate fits for the CAT and the HPX domains and the best fit orientation tensor values are shown in A and B respectively. The poor global fit assuming the rigid X-ray structure is shown in C, and that obtained by best fitting the reciprocal orientation of the two domains to the RDC values is shown in D. In all panels, triangles refer to the CAT and filled circles refer to the HPX domain.

obtained by the linker randomizations representing possible conformations of MMP-12 in solution. None of these random models yielded a computed scattering in agreement with the experimental data. This was not unexpected, as SAXS "sees" the conformational and orientational average over the large (approx. 10^{16}) ensemble of protein molecules in the illuminated specimen volume. A simple average intensity of the generated pool also did not agree with the experiment, suggesting a nonrandom configuration of the linker in MMP-12. To assess the preferable conformations in solution, the EOM method[79] was used. Given a representative pool of (random) structures, the method employs a genetic algorithm to select the ensembles from this pool that best fit the experimental data, as explained in materials and methods. Several EOM runs yielded reproducible ensembles neatly fitting the experimental data with discrepancy χ in the range 1.1−1.3, and a typical fit provided by the ensemble selected by EOM is given in Figure 5, curve 4.

All the fits from different EOM runs are graphically indistinguishable from curve 4 in Figure 5.

The R_g distributions of the particles in the initial pool and in the selected ensembles are compared in the insert to Figure 5. The former distribution is rather broad, and covers the R_g range from about 23 to 50 Å, corresponding to extremely compact and completely extended domain configurations, respectively. In contrast, the R_g distribution of the selected ensembles displays a relatively sharp peak around $R_g = 28-29$ Å, including about 55% of the particles in the selected ensembles. Visual inspection of the models in the peak indicates, not unexpectedly, that they have a shape similar to that of the MMP-12 structure in the crystal (although with varying interdomain orientations). In contrast, not a single structure with $R_g < 27$ Å was selected in multiple EOM reconstructions, indicating that models similar to the crystal structure of FL-MMP-1 were never present. These results indicate that the present crystal structure of FL-MMP-12 may be significantly present also in solution, but also that the protein experiences noticeable conformational flexibility, as revealed by the presence of a more or less uniform distribution of particles in the range between $R_g = 30$ and 50 Å and of a significant spike, which was always observed for the most extended particles (Figure 5, insert). We also tried to explore the possibility of a two-state exchange situation allowing for only two conformations in the mixture. The two-state fits were however always poorer than those from twenty-state EOM populations. In particular, by fixing the first state to be the crystal structure and allowing EOM to select the second state, discrepancies not better than 1.5−1.6 could be obtained. These results suggest that MMP12 adopts a manifold of conformations in solution, in full agreement with NMR observations.

Concluding Remarks and Biological Implications

The present data demonstrate that full-length MMP-12 shows relative mobility of its catalytic and hemopexin domains. The observation of R_1 and R_2 values intermediate between those of the isolated domains and those expected for any rigid structure of the full-length protein is particularly striking in this respect.

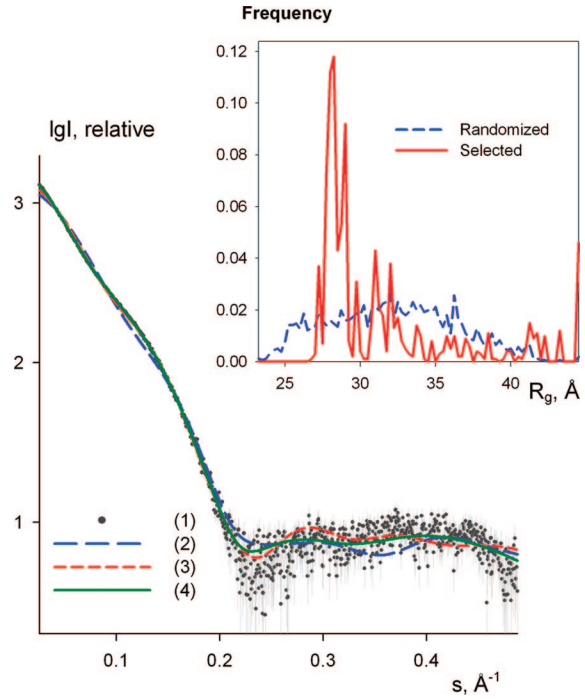


Figure 5. Experimental X-ray scattering from the NNGH-inhibited, cadmium(II)-substituted Phe171Asp/Glu219Ala double mutant of FL-MMP-12, and scattering from the models. (1) experimental data with error bars; (2−3) computed scattering from the crystallographic models of FL-MMP-1 and FL-MMP-12, respectively; (4) a typical fit by the selected ensemble of structures. The logarithm of the scattering intensity is plotted against the momentum transfer. Insert: the frequency of the models with the given R_g in the initial pool of structures with randomized interdomain linkers (blue broken line) and in the selected ensembles (red solid line); the latter distribution is obtained by the averaging of several EOM runs. Both R_g distributions are normalized to the integral value of unity.

Indeed, even in the case of calmodulin, the two-domain protein that constitutes a paradigmatic example of large interdomain mobility,[43,45,90,91] the R_1 and R_2 values are only modestly different from what expected for a rigid structure.[45,90] Apparently, in the case of FL-MMP-12 the reorientation of the backbone NH vectors with respect to the magnetic field occurs on a time scale that is faster than the rotational time of the whole molecule, whereas in calmodulin it is of the same order.[45,90] Conversely, the amplitude of the motion is probably lower for FL-MMP-12 than it is for calmodulin, as judged from the SAXS data that suggest that the molecules spend about half of the time in a conformation that is more or less as compact as the solid state structure. A similar behavior is probably experienced by the two-domain protein xylanase Cex, whose flexibility has been recently demonstrated.[89]

The present data are a significant example of the synergy between NMR and SAXS techniques.[92–96] NMR provides evidence of conformational freedom and of the time scale, whereas SAXS provides insight into the types and variety of the sampled conformations. About half of the conformations that are likely to be experienced by FL-MMP-12 in solution are as compact as the solid state structure but with different relative orientations of the two domains, while another half a re more extended, and some even highly extended. In this respect the X-ray structure does not provide meaningful information on the conformation of the protein in solution.

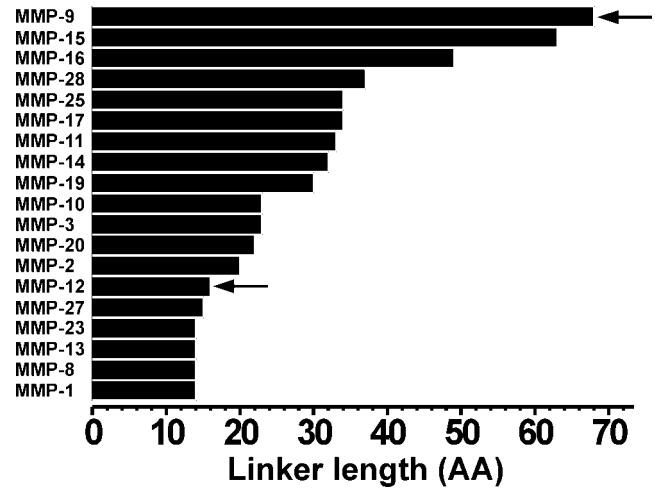


Figure 6. Linker lengths in matrix metalloproteinases. Arrows indicate MMPs for which interdomain mobility has been demonstrated (26 and present work).

Relative mobility of the CAT and HPX domains has been recently suggested for MMP-9, where the two domains are separated by the OG domain.[97] In the case of MMP-9 it has been argued[26] that the long and flexible OG domain may mediate protein−substrate interactions. Independent domain movements might even mediate enzyme translocation on a collagen fibril.[98,99] Another interesting possibility is that domain flexibility can mediate the activation of the enzyme and the cleavage of the pro-domain by promoting long-range conformational transitions induced by the binding of the activator proteins.[100] It is possible that MMP-12 can be representative of all other MMPs where the two domains are connected by a short linker rather than by a long one (Figure 6). However, this has to be demonstrated for those cases in which the contact area between the CAT and HPX domains is much larger than in the present case. Notably, the possibility of reorienting the hemopexin with respect to the catalytic domain during catalysis has been often invoked to rationalize the fact that, for collagenases (MMP-1, MMP8, and MMP-13), the catalytic domain alone is not able to attack collagen, whereas the full-length protein does.[18,19]

On the contrary, noncollagenase MMPs such as MMP-2 and MMP-12 do not degrade the native triple helix collagen. The

(90) Wang, T. Z.; Frederick, K. K.; Igumenova, T. I.; Wand, A. J.; Zuiderweg, E. R. P. *J. Am. Chem. Soc.* **2005**, *127*, 828–829.
(91) Bertini, I.; Gupta, Y. K.; Luchinat, C.; Parigi, G.; Peana, M.; Sgheri, L.; Yuan, J. *J. Am. Chem. Soc.* **2007**, *129*, 12786–12794.
(92) Mattinen, M. L.; Paakkonen, K.; Ikonen, T.; Craven, J.; Drakenberg, T.; Serimaa, R.; Waltho, J.; Annila, A. *Biophys. J.* **2002**, *83*, 1177–1183.
(93) Perry, A.; Tambyrajah, W.; Grossmann, J. G.; Lian, L. Y.; Scrutton, N. S. *Biochemistry* **2004**, *43*, 3167–3182.
(94) Grishaev, A.; Wu, J.; Trewhella, J.; Bax, A. *J. Am. Chem. Soc.* **2005**, *127*, 16621–16628.
(95) Marino, M.; Zou, P. J.; Svergun, D.; Garcia, P.; Edlich, C.; Simon, B.; Wilmanns, M.; Muhle-Gol, C.; Mayans, O. *Structure* **2006**, *14*, 1437–1447.
(96) Tsutakawa, S. E.; Hura, G. L.; Frankel, K. A.; Cooper, P. K.; Tainer, J. A. *J. Struct. Biol.* **2007**, *158*, 214–223.
(97) Van den Steen, P. E.; Van Aelst, I.; Hvidberg, V.; Piccard, H.; Fiten, P.; Jacobsen, C.; Moestrup, S. K.; Fry, S.; Royle, L.; Wormald, M. R.; Wallis, R.; Rudd, P. M.; Dwek, R. A.; Opdenakker, G. *J. Biol. Chem.* **2006**, *281*, 18626–18637.
(98) Overall, C. M.; Butler, G. S. *Structure* **2007**, *15*, 1159–1161.
(99) Saffarian, S.; Collier, I. E.; Marmer, B. L.; Elson, E. L.; Goldberg, G. *Science* **2004**, *306*, 108–111.
(100) Rosenblum, G.; Meroueh, S.; Toth, M.; Fisher, J. F.; Fridman, R.; Mobashery, S.; Sagi, I. *J. Am. Chem. Soc.* **2007**, *129*, 13566–13574.

real function of these noncollagenase proteins is still unknown. MMP-2 is able to degrade gelatin, a product of partial hydrolysis of collagen, and MMP-12 elastin, but we do not know if these are their real physiological roles. For MMP-12 it has been reported that the catalytic domain alone is able to degrade elastin even without the hemopexin domain. However, if during evolution the hemopexin domain has been maintained in spite of selective pressure, it is difficult to believe that it is useless. In this respect, the present discovery of relative mobility of the two domains in MMP-12 might be important for this still unknown function.

Acknowledgment. We are grateful to Prof. Irit Sagi and to Prof. Christopher M. Overall for the fruitful discussions. This work was supported by EC (Projects: MEST-CT-2004-504391, LSHG-CT-2004-512077, SPINE II LSHG-CT-2006-031220, ORTHO AND PARA WATER no 005032 and Nano4Drugs no LSHB-CT-2005-019102), by MIUR (PRIN 2005, Prot. N. 2005039878, Prot. RBLA032ZM7), and Ente Cassa di Risparmio di Firenze. We acknowledge the support and assistance of the DESY (Hamburg) synchrotron radiation facilities for the SAXS data collection.

Supporting Information Available: NMR chemical shift values of FL-MMP-12 and of the hemopexin domain of MMP-12. The list of upper experimental constraints used for structure calculations. RDC values of FL-MMP-12. Statistical analysis of the NMR structures of the hemopexin domain of MMP-12. Relaxation data. Data collection, processing statistics, and refinement statistics for the crystallographic data set; Figure S1 and S2. This material is available free of charge via the Internet at http://pubs.acs.org.

JA710491Y

5. Biotin-Tagged Probes for Profiling of MMPs Expression and Activation: Design, Synthesis and Binding Properties[1]

Elisa Dragoni, Vito Calderone, Marco Fragai, **Rahul Jaiswal,** Christina Nativi*, Claudio Luchinat

Submitted to *Bioconjugate Chemistry*

---

[1], C.L., C.N. & M.F. designed the research, R.J. done all the Molecular Biology work and sample preparation, E.D. done the organic synthesis work, V.C. done the X-ray crystallography experiment, M.F. done the NMR experiments.

Biotin-Tagged Probes for Profiling of MMPs Expression and Activation: Design, Synthesis and Binding Properties

Elisa Dragoni,[1,2] *Vito Calderone,*[1] *Marco Fragai* [1,3] *Rahul Jaiswal,*[1] *Claudio Luchinat,*[*,1,3] *Cristina Nativi*[*,1,2]

[1] Magnetic Resonance Center (CERM) – University of Florence, Via L. Sacconi 6, 50019 Sesto Fiorentino, Italy

[2] Department of Organic Chemistry, University of Florence, Via della Lastruccia 13, 50019 Sesto Fiorentino (Italy)

[3] Department of Agricultural Biotechnology, University of Florence, Via Maragliano, 75-77, 50144 Florence, Italy.

[*] Prof. Claudio Luchinat
Magnetic Resonance Center (CERM)
University of Florence
Via L. Sacconi 6
50019 Sesto Fiorentino, Italy
e-mail: luchinat@cerm.unifi.it;
Tel.: +390554574262
Fax: +390554574253

[*] Prof. Cristina Nativi
Department of Organic Chemistry
University of Florence
Via della Lastruccia 13
50019 Sesto Fiorentino, Italy
e-mail: cristina.nativi@unifi.it;
Tel.: +390554573540
Fax: +390554573531

Abstract

The design and synthesis of biotin chain-terminated inhibitor (BTI) showing high affinity for matrix metalloproteinases (MMPs) on one side and high affinity for avidin through the biotinylated tag on the other, is reported. The affinity of the designed BTI towards five different MMPs has been evaluated and the formation of a highly stable ternary system Avidin-BTI-MMP clearly assessed. This system will permit the development of new approaches to detect, quantify and collect MMPs in biological samples, with potential applications in vivo.

KEYWORDS: MMPs, Probe, Biotin, NMR, X-ray

Introduction

The design of molecular probes able to selectively bind biomolecular targets is gaining importance not only for their applications in molecular biology, but also for the development of new techniques for the early stage diagnosis and therapy of several pathologies.[1] Rational drug design driven by structural data is currently applied to speed-up the design of new candidate drugs in pharmaceutical research.[2-5] The same approach is even more important to design molecular probes where, besides classical requirements, additive structural and functional features are required.

Matrix metalloproteinases (MMPs), also known as matrixins, are proteinases that participate in extracellular matrix (ECM) degradation[6] and in several extracellular processes.[7-10] Under normal physiological conditions the activity of MMPs is strictly regulated by transcription/activation of zymogens precursors, interaction with specific ECM components and inhibition by endogenous inhibitors (TIMPs).[11-14] An imbalance of these activities may results in diseases such as atherosclerosis, thrombosis or heart failure.[15;16] As a matter of fact, expression and activation of MMPs are increased in these pathologies, as well as in almost all human cancers, compared to physiological conditions.[17]

MMPs have long been associated with cancer and metastasis. Indeed, ECMs degradation by MMPs does promote tumor invasiveness. Furthermore, it has been recently demonstrated that MMPs carry out functions other than promotion of invasion, and that they likely take part in the development of cancer even before invasion. Noteworthy, the degree of overexpression of MMP in cancer tissues is strictly associated with the level of tumor invasiveness and with its tendency to form metastasis. It is therefore apparent that the development of efficient molecular devices designed to reveal the overexpression of MMP may be a powerful tool for the early detection of tumors and the assessment of aggressiveness.[18]

Studies focused on the detection of overexpression of MMPs have already been reported[19] and probes to profile MMPs in tissues synthesized and tested.[20;21] In particular, hydrolizable, near-infrared fluorescent molecules have been successfully tested in mice to detect fibrosarcoma.[22] Unfortunately, this innovative technique suffers from the serious limitation represented by the limited penetration of light (a few millimetres); such a degree of penetration is only sufficient for the readout of superficial subcutaneous tumors (< 10 mm).

An interesting approach to detect overexpression of MMP in atherosclerosis has also been published. Low molecular weight inhibitors, able to interact with the active site of MMPs involved in atherosclerotic plaque disruption, have been targeted with nuclear agents and tested *in vivo*.[23-25] Although the results reported were extremely encouraging, the approach presented a scarce flexibility, due to experimental difficulties in preparing the targeted inhibitors and to limited availability of suitable nuclear agents.

A significant improvement in the development of MMPs-conjugated probes to detect pathological events may be the rational design of biotin-labeled inhibitors,[26] capable of binding to MMPs exclusively and with high affinity. These tagged molecules relay on the well known avidin-biotin molecular recognition system and could be used in conjunction to commercially available biotin-labelled probes (es. radionuclide chelates) to efficiently detect pathological tissues by imaging the local overexpression of MMPs. The use of such molecules is not limited to *in vivo* imaging but can be extended to *in vitro* applications such as MMPs purification by using avidin columns or devices.

Capitalizing on our expertise on expression and structural characterization of MMPs[27-29] and on the rational design and synthesis of MMP inhibitors,[30-35] in the present

paper we report the design and synthesis of, carboxylic-based, biotinylated inhibitors able to show high affinity for several MMPs and at the same time able to strongly interact with avidin through the biotinylated tag.

In clinical therapy, inhibitors' selectivity for a specific MMP is dramatically important, so much so that the lack of selectivity is believed to have been the main reason for failures of clinical trials in the past. On the contrary, for the molecules here proposed, selectivity is not an issue. Indeed, ideal probes to detect overexpression of MMP, must not be selective inhibitors vs. a specific MMP but they should instead selectively recognize MMPs among other proteins. These new biotinylated molecules, designed to bind selectively to MMPs and to be recognized by avidin, might therefore be successfully used in the visualization of pathological tissues and to evaluate the aggressiveness of the pathology by monitoring radioactivity. Moreover, it might be possible to test *in vivo* the efficacy of clinical treatments against a specific disease by just monitoring the level of MMPs produced during the therapy.

Results and Discussion

Aryl-sulfonamide derivatives have been largely exploited to synthesize broad-spectrum MMPs' inhibitors. These scaffolds present different positions which can be easily functionalized by introducing new functional groups. This feature and their easy synthetic accessibility make sulfonamide derivatives suitable candidates to design biotin-tagged probes for MMPs. As zinc binding groups (ZBG), the carboxylic acid was selected despite its relatively weak affinity for the metal ion.[36;37] The choice of a carboxylic acid to replace the most commonly used hydroxammic acid warrants a higher stability of the inhibitor under physiological conditions, avoiding the formation of undesired and toxic degradation products.

In all MMPs, the flexibility of the protein loops forming the active site makes these enzymes able to modify the shape of the catalytic pocket to accommodate ligands with different structures [38-41]. In particular, analyzing the structures available on protein data bank, it is apparent that the volume of the S_1' cavity can be changed to host ligands bearing lipophilic groups of different size. However in MMP-1 and MMP-7 the S_1' cavity is so small that bulky groups generally misfits. Therefore, in order to probe all members of MMPs' family, two scaffolds, bearing either a bulky biphenyl moiety or a smaller methoxyphenyl group, were considered as possible bases for the synthesis of BTIs.

Among all the scaffolds evaluated, molecules **1** and **2** (Figure 1) were selected for their low Ki values against five different MMPs (see Table 1). Noteworthy, the two scaffolds exhibit a sizable different selectivity and affinity for the MMPs chosen. In particular, inhibitor **1**, due to the presence of a biphenyl moiety on the sulfonamide sulfur showed inhibition constants in the nanomolar range. Only for MMP-1 and MMP-7 the affinity was in the low-micromolar range due to the relatively small size of the S_1' pocket.[42] On the contrary, **2** is a weaker MMP inhibitor with micromolar values of Ki for all the MMPs tested.

Matrix metalloproteinase 12 (MMP-12) performs proteolysis of several ECM components including elastin, laminin and type IV collagen. The stability of the active MMP-12's catalytic domain in the presence of weak inhibitors, the relatively large S_1' pocket, and the high yield of the ^{15}N-enriched protein in minimal medium, make this enzyme an ideal model to evaluate BTI candidates.

To explore the binding mode of the two scaffolds and to properly design the biotin-tag derivatives, the two inhibitors were soaked in crystals of MMP-12 and the structure of the complexes solved by X-ray crystallography. Analysis of the crystal structures (see Figure 2) revealed that the inhibitors share a similar binding mode, with the catalytic zinc ion coordinated by the oxygens of the free carboxylate, the sulfonyl oxygen H-bonded with the NH of the Ala 182, and the liphophilic groups nested inside the S1' pocket. The structural analysis accounts for the larger affinity of **1** for MMP-12 with respect to scaffold **2**, as observed in the enzymatic assays. Indeed, in the MMP-12 - **1** complex, larger hydrophobic interactions result from the larger surface-to-surface contact area due to the presence of the biphenyl moiety. More important for the biotin-tag inhibitors design was the solvent exposure of the sulfonamide nitrogen revealed by the crystal structures for both ligands. These structural data address the question related to the identification of the position where the tag can be inserted without spoiling the binding affinity.

A further critical point in developing efficient biotinylated probes for *in vivo* and *in vitro* applications is represented by the choice of a suitable linker able to prevent any steric clashes between avidin and the target biomolecule. While a too short spacer may impair the ability of the probe to bind both MMP and avidin, the use of a too long spacer could affect the bioavailability, stability and solubility of the whole molecule. In

particular, a too long linker can increase the susceptibility of the probe to be split by enzymes such as the biotinidase that can degrade the molecule in vivo.[43] Once more, the availability of the three-dimensional structure of both avidin (present in protein data bank, code: 2avi) and MMP-12's catalytic domain provided the background information needed for a rational molecular design of the biotin-tagged inhibitors we conceived.

A fourteen-membered spacer was designed to link the biotin residue to the carboxylic-methyl ester sulfonamide **3**.

The biotinylated inhibitor **4** was obtained by conjugating biotin aminoexylamide derivative **5** to the bromoexyl methylester **6**, through a secondary amino group (Scheme 1). Biotin was first reacted under known procedures[43] with the N-Boc-1,6-diaminohexane **7** to give the amido derivative **8** which, after removal of *tert*-butyloxy carbonyl protecting group, **11** was treated with the bromosulfonamide **6** in the presence of cesium hydroxide, in dimethylformamide as solvent, to give the enantiomerically pure ester **9** in 55% yield (see Supporting Information) (Scheme 2). Following the synthetic pathway reported for **4** the biotin-tagged MMP-12's inhibitor **10** was also prepared from **11** (Scheme 2) and bromosulfonamide **12** as diasteromerically pure compound (see Supporting Information) (Scheme 3).

The affinity properties of the two BTI derivatives toward the selected MMPs were firstly evaluated by fluorimetric assay. While compound **4** had lower Ki values as compared to the corresponding precursor **1**, BTI **10** exhibited comparable or weaker affinity with respect to its precursor **2**. The micromolar Ki for all the tested MMPs unambiguously demonstrated that **10** could not have been used as a probe to detect these enzymes and it was not further investigated. Conversely, BTI **4** exhibited a high affinity for all the considered MMPs but one (MMP-1). The low affinity for MMP-1, already predicted during the design of the molecule, prevent the use of **4** to probe this enzyme.

Several attempts aimed to obtain the X-ray structure of the biotin-tag inhibitors in complex with MMP-12, either by soaking or co-crystallization, failed, probably for steric reasons owing to the tag. Therefore, interaction of compound **1** and of its biotin-tag derivative **4** with the catalytic domain of MMP-12 was further investigated following the chemical shift perturbations on ^{15}N-^{1}H HSQC spectra using ^{15}N labeled protein samples. The chemical shift is a sensitive indicator of protein-ligand interactions and can be use to identify the binding site when the resonance assignment is available.[44] The complete

assignment of the catalytic domain of MMP-12 complexed with the weak inhibitor acetohydroxamic acid is available in our laboratory. Increasing amounts of **1** and **4** were added to 200 μM solutions of the protein up to an equimolar concentration. Although several NH signals were shifted by adding the inhibitors, as many as 134 and 129 cross-peaks belonging to protein adducts with **1** and **4,** respectively**,** where easily re-assigned. The intensity of signals belonging to NH groups perturbed by the addition of the ligands decreased linearly, while new cross-peaks corresponding to the BTI-bound form of the protein appeared and gradually increased their intensity, in agreement with the effect produced by a slow-exchanging ligand.

The binding mode of ligands **1** and **4**, respectively, was analyzed and compared by reporting Garrett values [$\Delta\delta$(NH)] on a surface representation of the MMP-12 structure (Figure 3). As it is evident from Figure 3, for both inhibitors, the region with significant ^{15}N-^{1}H HSQC spectral perturbations involves the catalytic pocket. The analysis rule of Garrett values out relevant additive interactions of the biotinylated-tail present in BTI **4** with different protein regions, suggesting a different origin for the observed improved affinity. A hypothetical binding mode with the biotinylated-tail of BTI **4** fitting in the S1' cavity can be completely excluded by the chemical shift perturbations on the ^{15}N-^{1}H HSQC spectra that are similar for **1** and **4** as well as by the Ki values of the BTI **10**, which, although biotinylated, showed a low affinity for the enzyme. Conversely, the effects monitored on the aminoacids forming the catalytic pocket account for a better fit of **4** into the active site of the enzyme that stabilizes the protein-ligand adduct.

Although a high affinity for the target protein is a basic requirement for the proposed bifunctional ligand to play its role, only the ability of the latter to bind MMP and avidin simultaneously could prove its real efficacy as effective BTI. Two different approaches based on NMR spectroscopy and on Surface Plasmon Resonance (SPR) respectively were followed to demonstrate the ternary Avidin-BTI-MMP interaction. Several NMR parameters are strongly influenced by the molecular weight and by the dynamical properties of the investigated macromolecules. Classical NMR experiments such as ^{15}N-^{1}H HSQC usually performed to investigate biomolecules up to 30-40 kDa are inefficient for larger systems due to the fast transverse relaxation rate that broaden the signals beyond the detection threshold.[45] Therefore, NH resonances can act as probes to investigate the formation of large protein-protein complexes.[46] In the present case, the

evolution of the NH signals of the ^{15}N MMP-12 catalytic domain complexed with the BTI **4** in ^{15}N-^{1}H HSQC spectra was exploited to monitor the interaction with non-labeled avidin, given that the binding of the biotin moiety to one of the four binding sites on the tetrameric avidin, should provide an 87 kDa. Since multiple binding for each avidin molecule may occur, even larger protein complexes are expected in solution. As shown in Figure 4 (panel D), the addition of an equimolar solution of the 65 kDa avidin to a solution of the ^{15}N labeled catalytic domain of MMP-12 complexed with the BTI **4** caused the complete disappearance of all NH cross peaks without any precipitation of the sample.

To exclude the presence of aspecific protein-protein interactions, avidin was added to a solution containing 200 μM of NNGH-inhibited-MMP-12 catalytic domain: in this case no changes were detected in the HSQC spectra map. Therefore, we can reasonably conclude that the strong protein-protein interaction, monitored by the ^{15}N-^{1}H HSQC experiments, is mediated by the BTI **4** which is able to bind MMP-12 and avidin and to prevent, at the same time, an extensive reciprocal reorientation of the complex subunits. Although NMR data allowed us to ascertain that **4** does bind both avidin and MMP-12 side-by-side, the micromolar concentrations used for these experiments do not provide information on how this simultaneous interaction may affect the binding affinity of the inhibitor to MMP-12. Indeed, the maintenance of a nanomolar affinity for MMP-12 of **4** when bound to avidin is a mandatory feature for a successful application of this BTI as probe to detect MMP's overexpression and activation, as well as to use the bifunctional linker for in vitro assays and in protein purification protocols.

Reasonable steric effects related to the presence of the hemopexin domain need to be evaluated to demonstrate the real binding capability of the probe toward functionally active enzymes.[47] In this respect, the availability of the active full length MMP-12 allowed us to ascertain the efficiency of the biotinylated inhibitor to bind simultaneously avidin and the physiologically active enzyme.

Surface Plasmon Resonance (SPR) is widely applied to investigate protein-protein and protein-ligand interactions. Information on affinity constant, on binding kinetic as well as on binding enthalpy can be quickly obtained by using small amounts of protein samples. In particular, the sensitivity of the SPR technique permits to investigate high affinity interactions and to monitor the formation of ternary complexes. SPR

analysis were carried out using a ProteOn XPR36 parallel array biosensor equipped with avidin sensor chips (NLC). This chip has an avidin-coated surface that can be used to immobilize biotinylated molecules and macromolecules.[48] In our study BTI **4** was immobilized on channels one to five of the chip by injecting PBS-T buffered solutions of the biotinylated molecule at different concentrations. The sixth channel was left as blank. The assay performed by injecting solutions of the full length MMP-12 on the chip channels and fitting the sensorgrams at four different concentrations. The measured K_D (4.6 ± 0.3 nM) was in excellent agreement with the fluorimetric assay.

Conclusion

In summary, the present structure-based design led to the development of a potent biotin-probe to bind matrix metalloproteinases for possible application *in vitro* and *in vivo*. The structural requirements defined through the analysis of the X-ray structures allowed selecting the position on the scaffold where the linker could be introduced and the correct length of the biotinylated tag. As defined by NMR analysis, the origin of the larger affinity observed for BTI **4** is a better fit of the selected scaffold in catalytic pocket rather than additive interactions of the linker on the protein surface. The enzymatic assay demonstrated a nanomolar affinity of **4** not only for MMP-12 but also for MMP-7,-8,-13. Furthermore, the evolution of the ^{15}N-^{1}H HSQC spectra upon the addition of unlabelled avidin to solution of ^{15}N MMP-12 clearly showed that the fourteen-member spacer used to link the biotin unit to carboxylic arylsulfonamide inhibitor **1** and **2** to the biotin moiety ensures the simultaneous binding of the two proteins without permitting large interprotein mobility. SPR analyses clearly showed the ability of **4** to bind the functional active enzyme MMP-12 and avidin side-by-side. This simultaneous interaction did not affect negatively the nanomolar affinity of inhibitor **4** for MMP-12, given that the SPR data nicely reproduced the fluorimetric values.

ACKNOWLEDGMENT. This work was supported by EC (Projects: MEST-CT-2004-504391, Integrated Project - SPINE2-COMPLEXES n° 031220**,** Nano4Drugs n° LSHB-CT-2005-019102; SFMET n. 201640), by MIUR (FIRB Prot. RBLA032ZM7; FIRB Prot. RBIP06LSS2), Ente Cassa di Risparmio di Firenze, Genexpress.

Captions to the figures

Figure 1. Structure of the selected scaffolds **1** and **2**

Figure 2. X-ray structure of the catalytic domain of MMP-12 complexed with 1 (A) and with 2 (B). Only the ligand-binding site is shown for clarity.

Figure 3. Surface representation of the residues of the catalytic domain of MMP-12 affected by chemical shift perturbation upon the addition of **1** (A) and of **4** (B)**.** Areas of interaction with significant ^{15}N-^{1}H HSQC spectra perturbations are colored-coded according to the Garret values. Magenta $\Delta\delta(NH)>0.1$ ppm; Yellow $0.1>\Delta\delta(NH)>0.05$; Cyan $0.05>\Delta\delta(NH)>0.03$.

Figure 4. ^{15}N-^{1}H HSQC spectrum of MMP-12 catalytic domain in absence (A) and in presence (B) of compound **1** (equimolar concentration). The two spectra are superimposed to highlight the cross-peak shifts (C). The ternary complex formation Avidin-BTI-MMP is seen in panel D where most of the cross peaks disappeared

Captions to the schemes

Scheme 1. Retrosynthesis of BTI **4**

Scheme 2. Synthesis of BTI **4**

Scheme 3. Retrosynthesis of BTI **10**

Table 1. Inhibition constants (K_i)of the considered compounds toward the considered MMPs.

Compound	MMP-1	MMP-7	MMP-8	MMP-12	MMP-13
1	18 ± 2 µM	56 ± 6 µM	31 ± 4 nM	25 ± 2 nM	108 ± 12 nM
2	302 ± 25 µM	80 ± 9 µM	4.6 ± 5 µM	1.4 ± 0.2 µM	2.6 ± 0.3 µM
4	25 ± 3 µM	450 ± 50 nM	4 ± 0.5 nM	5 ± 0.6 nM	0.7± 0.2 nM
10	295 ± 40 µM	73 ± 7 µM	8 ± 1 µM	1.6 ± 0.2 µM	1.2 ± 0.1 µM

REFERENCES

(1) Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. *Science* **2006**, *312*, 217-224.

(2) Blundell, T. L.; Jhoti, H.; Abell, C. *Nat.Rev.Drug.Discov.* **2002**, *1*, 45-54.

(3) Johnson S.L.; Pellecchia M. *Curr.Top.Med.Chem.* **2006**, *6*, 317-329.

(4) Heckmann, D.; Meyer, A.; Marinelli, L.; Zahn, G.; Stragies, R.; Kessler, H. *Angew.Chem.Int.Ed.* **2007**, *46*, 3571-3574.

(5) Jahnke, W.; Blommers, M. J.; Fernandez, C.; Zwingelstein, C.; Amstutz, R. *Chembiochem.* **2005**, *6*, 1607-1610.

(6) Page-McCaw, A.; Ewald, A. J.; Werb, Z. *Nat.Rev.Mol.Cell Biol.* **2007**, *8*, 221-233.

(7) Parks, W. C.; Wilson, C. L.; Lopez-Boado, Y. S. *Nat.Rev.Immunol.* **2004**, *4*, 617-629.

(8) D'Alessio, S.; Fibbi, G.; Cinelli, M.; Guiducci, S.; Del Rosso, A.; Margheri, F.; Serrati, S.; Pucci, M.; Kahaleh, B.; Fan, P. S.; Annunziato, F.; Cosmi, L.; Liotta, F.; Matucci-Cerinic, M.; Del Rosso, M. *Arthritis Rheum.* **2004**, *50*, 3275-3285.

(9) Boire, A.; Covic, L.; Agarwal, A.; Jacques, S.; Sherifl, S.; Kuliopulos, A. *Cell* **2005**, *120*, 303-313.

(10) Fragai, M.; Nesi, A. *ChemBioChem* **2007**, *8*, 1367-1369.

(11) Knauper, V.; Will, H.; Lopez-Otin, C.; Smith, B.; Atkinson, S. J.; Stanton, H.; Hembry, R. M.; Murphy, G. *J.Biol.Chem.* **1996**, *271*, 17124-17131.

(12) Morrison, C. J.; Overall, C. M. *J.Biol.Chem.* **2006**, *281*, 26528-26539.

(13) Ra, H. J.; Parks, W. C. *Matrix Biol.* **2007**, *26*, 587-596.

(14) Overall, C. M.; Tam, E.; McQuibban, G. A.; Morrison, C.; Wallon, U. M.; Bigg, H. F.; King, A. E.; Roberts, C. R. *J.Biol.Chem.* **2000**, *275*, 39497-39506.

(15) Fingleton, B. *Curr.Pharm.Des.* **2007**, *13*, 333-346.

(16) Skiles, J. W.; Gonnella, N. C.; Jeng, A. Y. *Curr Med Chem* **2004**, *11*, 2911-2977.

(17) Egeblad, M.; Werb, Z. *Nat.Rev.Cancer* **2002**, *2*, 161-174.

(18) Bremer, C.; Tung, C. H.; Bogdanov, A.; Weissleder, R. *Radiology* **2002**, *222*, 814-818.

(19) Lepage, M.; Dow, W. C.; Melchior, M.; You, Y.; Fingleton, B.; Quarles, C. C.; Pepin, C.; Gore, J. C.; Matrisian, L. M.; McIntyre, J. O. *Mol.Imag.* **2007**, *6*, 393-403.

(20) David, A.; Steer, D.; Bregant, S.; Devel, L.; Makaritis, A.; Beau, F.; Yiotakis, A.; Dive, V. *Angew.Chem.Int.Ed.* **2007**, *46*, 3275-3277.

(21) Saghatelian, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt, B. F. *Proc.Natl.Acad.Sci.USA* **2004**, *101*, 10000-10005.

(22) Bremer, C.; Tung, C. H.; Weissleder, R. *Nat.Med.* **2001**, *7*, 743-748.

(23) Faust, A.; Waschkau, B.; Waldeck, J.; Holtke, C.; Breyholz, H. J.; Wagner, S.; Kopka, K.; Heindel, W.; Schafers, M.; Bremer, C. *Bioconjug.Chem.* **2008**, *19*, 1001-1008.

(24) Wagner, S.; Breyholz, H. J.; Law, M. P.; Faust, A.; Holtke, C.; Schroer, S.; Haufe, G.; Levkau, B.; Schober, O.; Schafers, M.; Kopka, K. *J.Med.Chem.* **2007**, *50*, 5752-5764.

(25) Wagner, S.; Breyholz, H. J.; Faust, A.; Holtke, C.; Levkau, B.; Schober, O.; Schafers, M.; Kopka, K. *Curr.Med.Chem.* **2006**, *13*, 2819-2838.

(26) Uttamchandani, M.; Wang, J.; Li, J. Q.; Hu, M. Y.; Sun, H. Y.; Chen, K. Y. T.; Liu, K.; Yao, S. Q. *J.Am.Chem.Soc.* **2007**, *129*, 7848-7858.

(27) Bertini, I.; Calderone, V.; Fragai, M.; Luchinat, C.; Maletta, M.; Yeo, K. J. *Angew.Chem.Int.Ed.* **2006**, *45*, 7952-7955.

(28) Bertini, I.; Calderone, V.; Fragai, M.; Luchinat, C.; Mangani, S.; Terni, B. *Angew.Chem.Int.Ed.* **2003**, *42*, 2673-2676.

(29) Bertini, I.; Calderone, V.; Fragai, M.; Luchinat, C.; Mangani, S.; Terni, B. *J.Mol.Biol.* **2004**, *336*, 707-716.

(30) Bertini, I.; Calderone, V.; Fragai, M.; Giachetti, A.; Loconte, M.; Luchinat, C.; Maletta, M.; Nativi, C.; Yeo, K. J. *J.Am.Chem.Soc.* **2007**, *129*, 2466-2475.

(31) Bertini, I.; Fragai, M.; Lee, Y.-M.; Luchinat, C.; Terni, B. *Angew.Chem.Int.Ed.* **2004**, *43*, 2254-2256.

(32) Calderone, V.; Fragai, M.; Luchinat, C.; Nativi, C.; Richichi, B.; Roelens, S. *ChemMedChem* **2006**, *1*, 598-601.

(33) Bertini, I.; Fragai, M.; Giachetti, A.; Luchinat, C.; Maletta, M.; Parigi, G.; Yeo, K. J. *J.Med.Chem.* **2005**, *48*, 7544-7559.

(34) Fragai, M.; Nativi, C.; Richichi, B.; Venturi, C. *ChemBioChem* **2005**.

(35) Mannino, C.; Nievo, M.; Machetti, F.; Papakyriakou, A.; Fragai, M.; Guarna, A. *Bioorg.Med.Chem.* **2006**, *14*, 7392-7403.

(36) Agrawal, A.; Romero-Perez, D.; Jacobsen, J. A.; Villarreal, F. J.; Cohen, S. M. *ChemMedChem* **2008**, *3*, 812-820.

(37) Puerta, D. T.; Cohen, S. M. *Curr.Top.Med.Chem.* **2004**, *4*, 1551-1573.

(38) Bertini, I.; Calderone, V.; Cosenza, M.; Fragai, M.; Lee, Y.-M.; Luchinat, C.; Mangani, S.; Terni, B.; Turano, P. *Proc.Natl.Acad.Sci.USA* **2005**, *102*, 5334-5339.

(39) Bhaskaran, R.; Palmier, M. O.; Bagegni, N. A.; Liang, X. Y.; Van Doren, S. R. *J.Mol.Biol.* **2007**, *374*, 1333-1344.

(40) Devel, L.; Rogakos, V.; David, A.; Makaritis, A.; Beau, F.; Cuniasse, P.; Yiotakis, A.; Dive, V. *J.Biol.Chem.* **2006**, *281*, 11152-11160.

(41) Tochowicz, A.; Maskos, K.; Huber, R.; Oltenfreiter, R.; Dive, V.; Yiotakis, A.; Zanda, M.; Bode, W.; Goettig, P. *J.Mol.Biol.* **2007**, *371*, 989-1006.

(42) Lovejoy, B.; Welch, A. R.; Carr, S.; Luong, C.; Broka, C.; Hendricks, R. T.; Campbell, J. A.; Walker, K. A. M.; Martin, R.; Van Wart, H.; Browner, M. F. *Nat.Struct.Biol.* **1999**, *6*, 217-221.

(43) Sabatino, G.; Chinol, M.; Paganelli, G.; Papi, S.; Chelli, M.; Leone, G.; Papini, A. M.; De Luca, A.; Ginanneschi, M. *J.Med.Chem.* **2003**, *46*, 3170-3173.

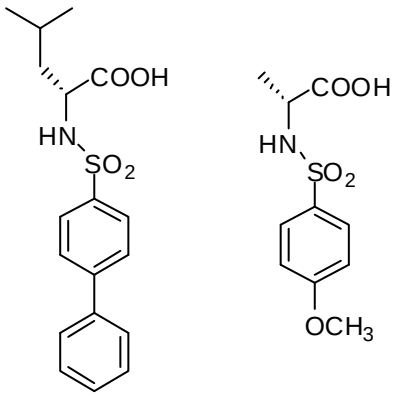
(44) Canales-Mayordomo, A.; Fayos, R.; Angulo, J.; Ojeda, R.; Martin-Pastor, M.; Nieto, P. M.; Martin-Lomas, M.; Lozano, R.; Gimenez-Gallego, G.; Jimenez-Barbero, J. *J.Biomol.NMR* **2006**, *35*, 225-239.

(45) Pellecchia, M.; Sebbel, P.; Hermanns, U.; Wuthrich, K.; Glockshuber, R. *Nat.Struct.Biol.* **1999**, *6*, 336-339.

(46) D'Silva, L.; Ozdowy, P.; Krajewski, M.; Rothweiler, U.; Singh, M.; Holak, T. A. *J.Am.Chem.Soc.* **2005**, *127*, 13220-13226.

(47) Bertini, I.; Calderone, V.; Fragai, M.; Jaiswal, R.; Luchinat, C.; Melikian, M.; Mylonas, E.; Svergun, D. I. *J.Am.Chem.Soc.* **2008**, *130*, 7011-7021.

(48) Bravman, T.; Bronner, V.; Lavie, K.; Notcovich, A.; Papalia, G. A.; Myszka, D. G. *Anal.Biochem.* **2006**, *358*, 281-288.



1 **2**

Figure 1

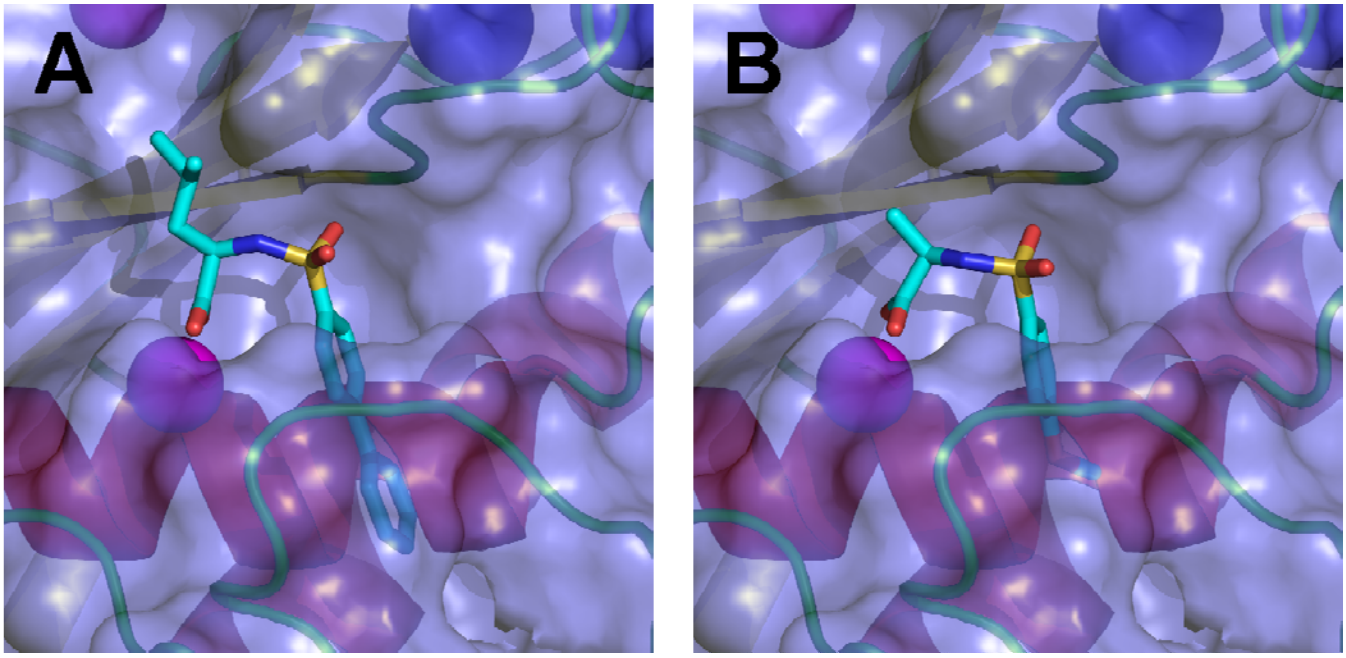


Figure 2.

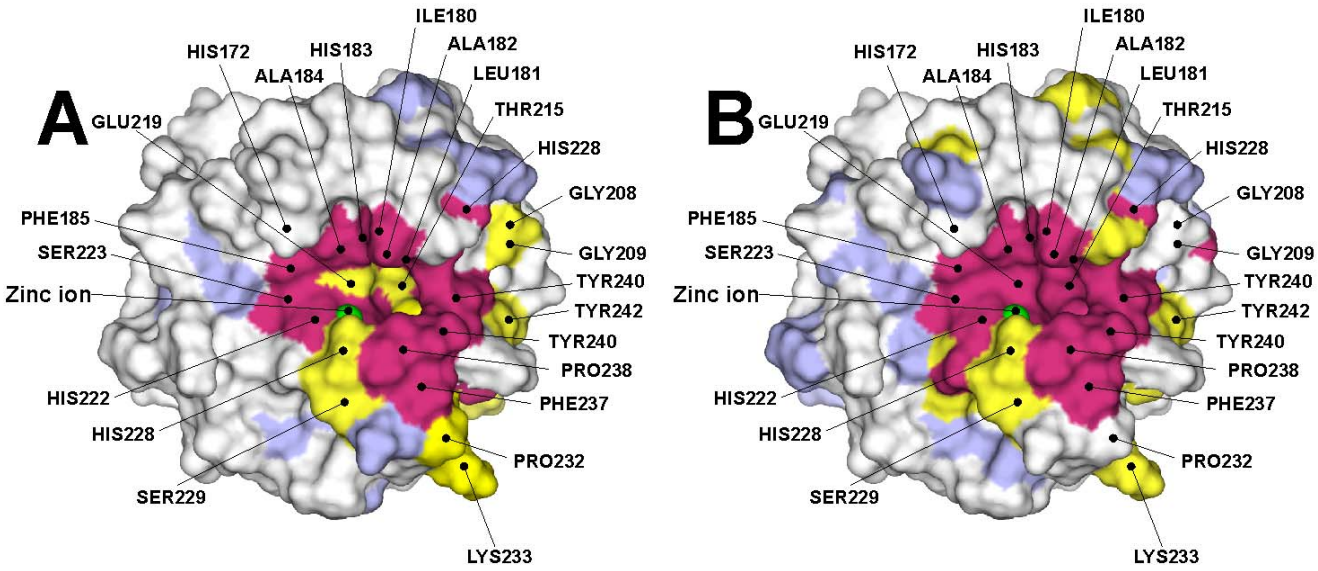

A
ILE180
HIS172
HIS183
ALA182
ALA184
LEU181
GLU219
THR215
HIS228
GLY208
PHE185
SER223
GLY209
Zinc ion
TYR240
TYR242
TYR240
PRO238
HIS222
PHE237
HIS228
PRO232
SER229
LYS233
B
ILE180
HIS172
HIS183
ALA182
ALA184
LEU181
GLU219
THR215
HIS228
GLY208
PHE185
SER223
GLY209
Zinc ion
TYR240
TYR242
TYR240
PRO238
HIS222
PHE237
HIS228
PRO232
SER229
LYS233

Figure 3.

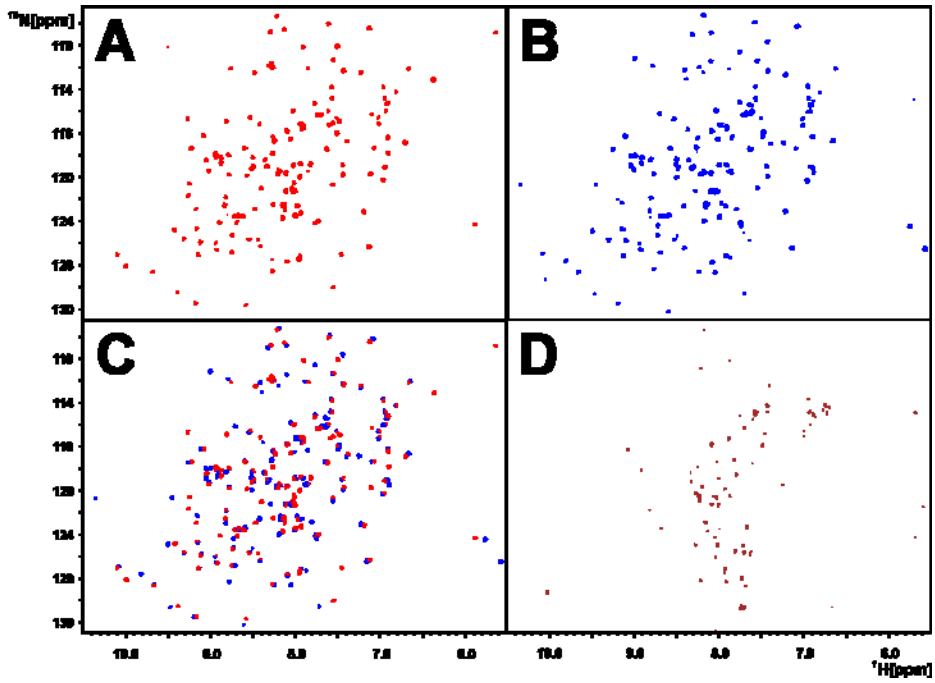
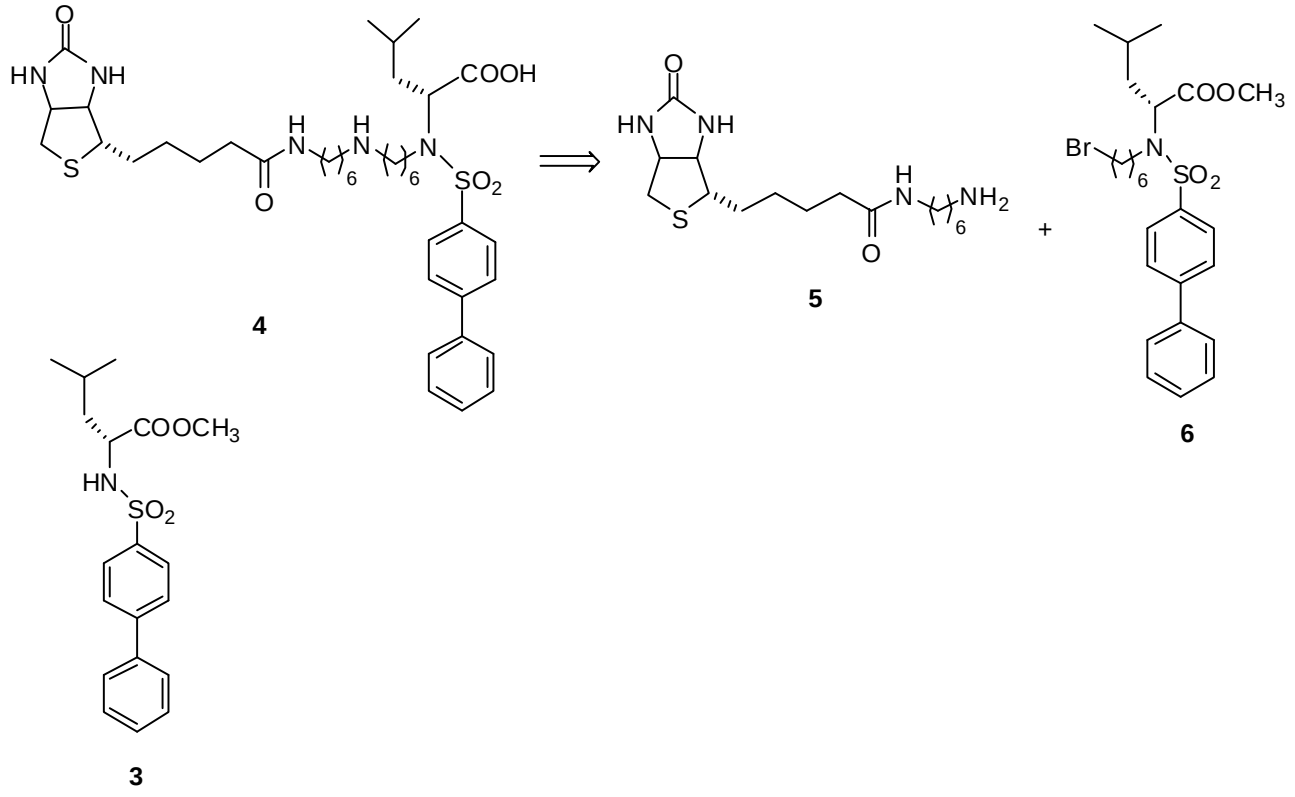
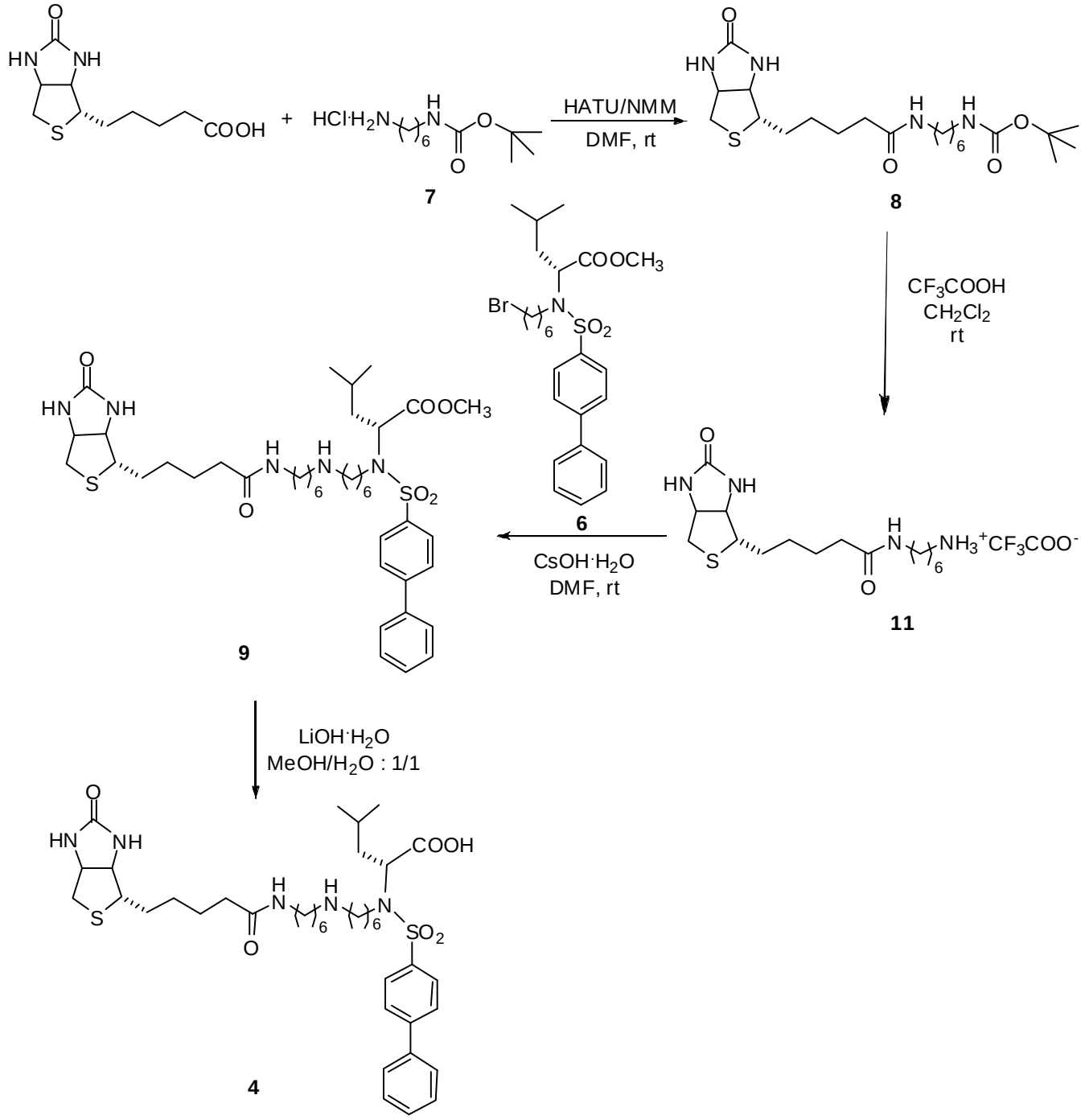


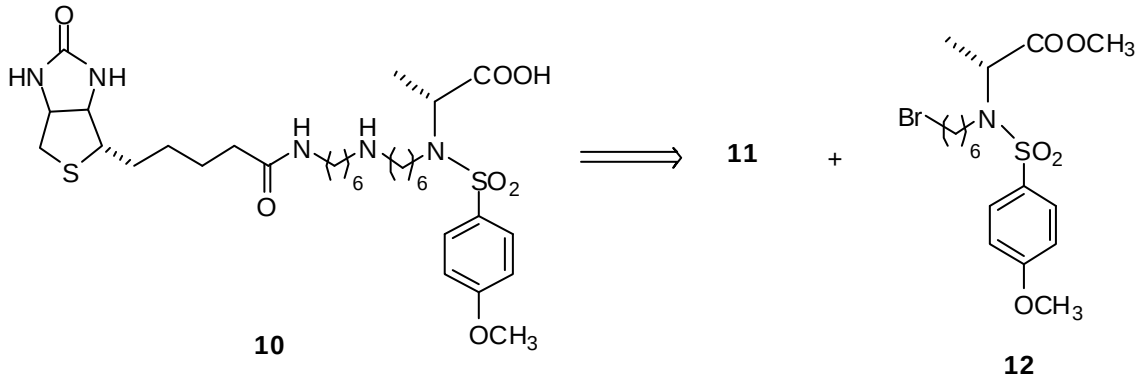
Figure 4.



Scheme 1.


HATU/NMM
DMF, rt
7
8
CF3COOH
CH2Cl2
rt
6
CsOH·H2O
DMF, rt
11
9
LiOH·H2O
MeOH/H2O : 1/1
4

Scheme 2.



Scheme 3.

6. Structural and thermodynamic analysis of a selective ligand for matrix metalloelastase[1]

[in preparation]

R.J. done all the Molecular Biology work and sample preparation.

Structural and thermodynamic analysis of a selective ligand for matrix metalloelastase

INTRODUCTION:

The selectivity is a crucial property for a drug, since side effects often arise from the interaction with proteins different from the target. When the target is a member of a high homology protein family, the availability of highly selective inhibitors is even more important. The matrix metalloproteinases (MMPs) is a family of structurally related extracellular zinc endopeptidases responsible for the degradation of several extracellular matrix proteins [(1)]. They play a critical role in both normal and pathological tissue remodeling processes [(2)]. Their uncontrolled proteolytic activity has been associated with a variety of diseases, including cancer, arthritis, multiple sclerosis, and atherosclerosis [(3)]. Therefore, there is substantial interest in developing MMP synthetic inhibitors for a variety of therapeutic indications. Results of the first clinical trials with broad spectrum MMP inhibitors in cancer therapy were disappointing, highlighting the need for better understanding of the exact role of each MMP during the different stages of tumor progression [(4)]. Recent research in this field has focused on the development of inhibitors that fully differentiate one MMP from another. This is a particularly difficult task, since the topology and nature of the residues in the enzyme's active site are highly conserved among the different MMPs. Moreover, parts of the MMP catalytic domain, which play a critical role in enzyme specificity, seem to be highly flexible [(5)]. This situation may explain why most previously reported MMP synthetic inhibitors preferentially inhibit some MMPs but are not exclusive inhibitors of a single MMP. Selective inhibitors for MMP-2, MMP-9 and MMP-12 have been reported recently [(6)]. In particular, a phosphinic peptide derivative has been shown to be a highly selective inhibitor of MMP-12 [(7)]. In such molecule an isoxazole ring placed in position P_1', is used as a rigid scaffold to project in the right orientation a biphenyl moiety able to fit the S_1' subsite of MMP-12. The two glutamates placed, respectively, in position P_2' and P_3' had been supposed to provide relevant interactions with aminoacids forming the catalytic pocket [(8)]. The relevance of MMP-12 as potential drug target was the prompt to investigate the origin of the observed selectivity with the aim of designing new inhibitors. Site directed mutagenesis experiments have been carried out to establish the real role of Lys177 and Thr239 in ligand binding and selectivity. This biochemical information have been

integrated and supported by a structural analysis at the solid state and in solution. In particular, the three-dimensional structure of the protein-ligand complex has been solved by X-ray crystallography and the interaction in solution analyzed by NMR. Crystal structures obtained by soaking and co-crystallization have been compared in order to explore possible different binding modes. The chemical shift mapping has been exploited to reveal long distance perturbation effects of the ligand on the protein resonances and to evaluate the occurrence of small, but energetically relevant conformational rearrangements of protein regions [(9)]. However, the structural data alone can not offer valuable insight into ligand selectivity since they do not provide information on the thermodynamics of the protein-ligand interaction. The analysis performed by Isothermal titration calorimetry (ITC) has revealed the thermodynamic driving force responsible for the picomolar affinity of the considered inhibitor allowing to new structure/energy correlations and toward new hypothesis on the origin of the observed selectivity.

MATERIALS AND METHODS:

Synthesis, characterization and inhibitory assays of the phosphinic inhibitor 1 toward MMP-12 have been already described [(10)].

Expression and purification of human MMP12 catalytic domain: The cDNA, encoding the fragment Gly106-Gly263 of the macrophage metalloelastase, was cloned into the pET21 vector (Novagen) using *NdeI* and *BamHI* as restriction enzymes and then transfected into *E. coli* strain BL21 Codon Plus cells. Uniform ^{15}N labelled protein was expressed by induction with 0.5 mM of IPTG at 37 C° for 4 h in M9 minimal media containing 15 mM of $(^{15}NH_4)_2SO_4$ and ^{13}C enriched glucose (Cambridge Isotope Laboratories). The inclusion bodies, containing the protein, were solubilized in a buffered solution with 20 mM Tris-HCl and 8 M urea at pH 8. The protein was then purified with a size exclusion chromatography (Pharmacia HiLoad Superdex 75 16/60) in 6 M urea and 50 mM sodium acetate. A second step of purification was performed on cation exchange column Mono-S (Pharmacia) using a linear gradient of NaCl up to 0.5 M.

The protein was refolded by using a multi-step dialysis against solution containing 50 mM Tris-HCl (pH=7.2), 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl and decreasing concentration of urea (from 4 M up to 2 M). The last two dialysis were performed against

solution containing 20 mM Tris-HCl (pH=7.2), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl and 200 mM of hydroxamic acid (AHA).

Fluorimetric assays. The inhibition constants for the compounds here investigated were determined evaluating their ability to prevent the hydrolysis of the fluorescent-quenched peptide substrate Mca-Pro-Leu-Gly-Leu-Dpa-Ala-Arg-NH_2 (Biomol Inc.). All measurements were performed in 50 mM HEPES buffer, with 10 mM $CaCl_2$, 0.05% Brij-35, 0.1 mM $ZnCl_2$ (pH 7.0), using 1 nM of enzyme and 1 μM of peptide at 298 K.

NMR measurements and assignment of the protein backbone

The experiments for the assignment of the protein backbone where performed on protein sample at a concentration of 0.7 mM (pH 7.2).

All NMR experiments were performed at 298 K and acquired on Bruker AVANCE 900, equipped with triple resonance CRYO-probes.

All spectra were processed with the Bruker TOPSPIN software packages and analyzed by the program CARA (Computer Aided Resonance Assignment, ETH Zürich) [(11)].

The backbone resonance assignment was obtained by the analysis of HNCA, HNCO, HNCACB and CBCA(CO)NH spectra performed at 900 MHz.

Crystallization, Data Collection and Resolution of the Crystal Structures: Crystals of human MMP12, already containing AHA from the refolding process, grew at 293 K from a 0.1 M Tris-HCl, 30% PEG 8000, 200 mM AHA, 1.0 M $LiCl_2$ solution at pH 8.0 using the vapor diffusion technique. The final protein concentration was about 10 mg/ml.

The complexes were obtained through soaking MMP12-AHA crystals with a solution containing the inhibitor itself in the presence of $LiCl_2$.

For the co-crystallized MMP12–Phosphinic adduct, 3 mM of NNGH was added to protein solution before mixing with a reservoir buffer containing 0.1 M Tris_HCl, 20% PEG 6000, 200 mM AHA, and 1.0 M LiCl2 at pH 8.0.

The data were measured in-house, using a PX-Ultra copper sealed tube source (Oxford Diffraction).

All the datasets were collected at 100 K and the crystals used for data collection were cryo-cooled without any cryo-protectant treatment. The crystals of all complexes had a mosaicity ranging from 0.3° to 0.8° and diffracted to a maximum resolution of 1.1 Å.

All the soaked adduct crystallize in the C_2 space group with one molecule in the asymmetric unit while co-crystallized complex of inhibitor crystallizes in $P2_12_12$ space group with one molecule in the asymmetric unit.

The data were processed in all cases using the program MOSFLM [12] and scaled using the program SCALA [13] with the TAILS and SECONDARY corrections on (the latter restrained with a TIE SURFACE command) to achieve an empirical absorption correction.

The structure of adduct with inhibitor was previously solved [14] using the molecular replacement technique; the model used was that of a molecule of human MMP12 (1OS9) while the structure of all other adducts were solved using the MMP12-AHA adduct (1Y93) as the model from where the inhibitor, all the water molecules and ions were omitted. The correct orientation and translation of the molecule within the crystallographic unit cell was determined with standard Patterson search techniques [15, 16] as implemented in the program MOLREP [17, 18]. The refinement was carried out using REFMAC5 [19] and for the atomic resolution datasets anisotropic B-factors were also refined. In between the refinement cycles the models were subjected to manual rebuilding by using XtalView [20]. The same program has been used to model all inhibitors. Water molecules have been added by using the standard procedures within the ARP/wARP suite [21] and for the atomic resolution datasets hydrogens were added at the riding positions and refined.

The stereochemical quality of the refined models was assessed using the program Procheck [22]. The Ramachandran plot for all structures is of very good quality.

The coordinates for all adducts are under deposition at the ProteinDataBank.

Calorimetry: The thermodynamic characterization of the interaction between the phosphinic inhibitor and the catalytic domain of MMP-12 has been carried out by performing ITC experiments at 298 K with a VP-ITC instrument (MicroCal, Inc., Northampton, MA). After an initial injection of 1 μL of 136 μM inhibitor, subsequent aliquots of 10 μL were stepwise injected into the sample cell containing a 12 μM solution of MMP12 catalytic domain (G106-G263), until complete saturation. All experiments were performed in 20 mM Tris (pH 7.2), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.1 M NaCl. Heats of dilution were measured by injecting the ligand into buffer and then subtracting the obtained values from the heats of binding. The thermodynamic parameters and K_A values were calculated fitting data to a single binding site model with ORIGIN 7.0 software (Microcal, Inc.) [23].

RESULTS AND DISCUSSION:

Site-directed mutagenesis studies: Several hypothesis on the structural determinants responsible for the tight binding and the high selectivity of phosphinic peptide inhibitor toward MMP-12, were based on a molecular model, generated starting from the x-ray data on a structurally related phosphinic inhibitor bound to the catalytic domain of MMP-11. The model revealed a close proximity of Glu in position P_2' and Glu in P_3' to Thr239 and to Lys177 respectively, suggesting a role of these two residues in ligand binding and in selectivity. To prove or disprove this hypothesis, a mutant of the catalytic domain of MMP-12, where Thr239 and to Lys177 were replaced by two glycine, has been expressed and the affinity for the phosphinic inhibitor determined by a fluorimetric assay. The K_i value was still in the picomolar range and in good agreement with that obtained for the wild-type protein. The retained affinity of the inhibitor for the mutant demonstrates that the interactions suggested by the model are too weak or even completely absent. The failure of the model in predicting the protein-ligand interactions was the prompt for a structural characterization of the complex.

Crystal structures In the crystal structures of the soaked and co-crystallized complex the inhibitor shows a similar but not identical binding pose.

Given the active site topology of MMPs, most inhibitors share similar features, i.e., the ability to bind to the metal ion, to the hydrophobic pocket termed S_1', and to the substrate binding groove. The phosphinic moiety is targeted to bind the catalytic zinc ion through one of the two oxygen atoms at a distance of 1.85 Å; the coordination environment for Zn is completed by His218, His222 and His228. The aromatic ring directly bound to the phosphinic group is involved in a stacking interaction with the ring of His222. The bromine atom bound to the latter ring is at 2.7 Å distance from the oxygen atom of Ala184 backbone. The bulky polycyclic aromatic part is targeted to deeply enter into hydrophobic S_1' pocket, to remove several waters molecules filling the cavity and to establish hydrophobic interactions.

The peptidic part shows significant H-bond interactions with the residues flanking the binding groove: four such interactions are present with the backbone atoms of Gly179, Leu181, Pro238 and Tyr240 with distances ranging from 2.5 Å to 3.3 Å.

A close up of the arrangement of the ligand in the active site of MMP12 is shown in Figure 1. As expected, the inhibitor share the same binding mode of other inhibitors

known in the literature and all relevant interactions (with the S_1' pocket, the metal, and the substrate binding groove) are in place. The comparison of the crystal structure of MMP-12-phosphinic complex with the free form of the protein has provided an insight into the effect of ligand binding on protein flexibility. The least-squares superposition between Free-MMP12 (2OXU) and MMP12–Phosphinic yields a global BB RMSD of 0.47 Å with only two narrow regions in which the deviation is significantly high: region 154–155 where it goes up to 1.7 Å, and region 230–233 where it goes up to 1.4 Å (Fig. 1*B*). Also the comparison of MMP12–Phosphinic with the NNGH inhibited structure (PDB code 1RMZ) has provided similar results, with a global BB RMSD of 0.41 Å and with major deviations localized in the same regions. Interesting, it is also the comparison between the soaked and co-crystallized structure of the MMP-12-Phosphinic inhibitor. For these two structure, the least-squares superposition yields a global BB RMSD of 0.53 Å, with large deviation again restricted to the same backbone segments. Collectively, these data provide an assessment of the conformational heterogeneity, previously reported for the same regions of the catalytic domain of MMPs, but more important, suggest that the mobility in these protein areas in not reduced by the interaction with the phosphinic inhibitor.

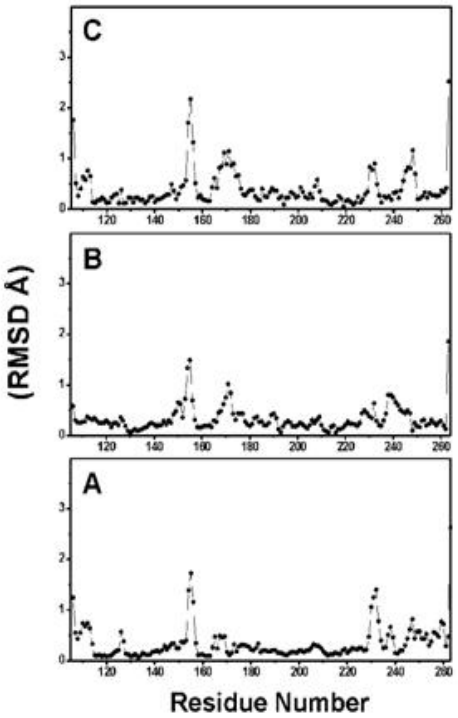


Figure 1. Global Backbone RMSD values comparison of (A) MMP12-Phosphinic Inhibitor, (B) MMP12-Phosphinic Inhibitor in presence of NNGH, (C) MMP12 soaked and co-crystallized with Phosphinic Inhibitor

Since chemical shifts are very sensitive to small changes in intermolecular interactions, the analysis of the Garret values [Δδ(NH)] provides a detailed map of the protein residues affected by the interaction upon the addition of the phosphinic inhibitor to the protein containing the weak inhibitor AHA.

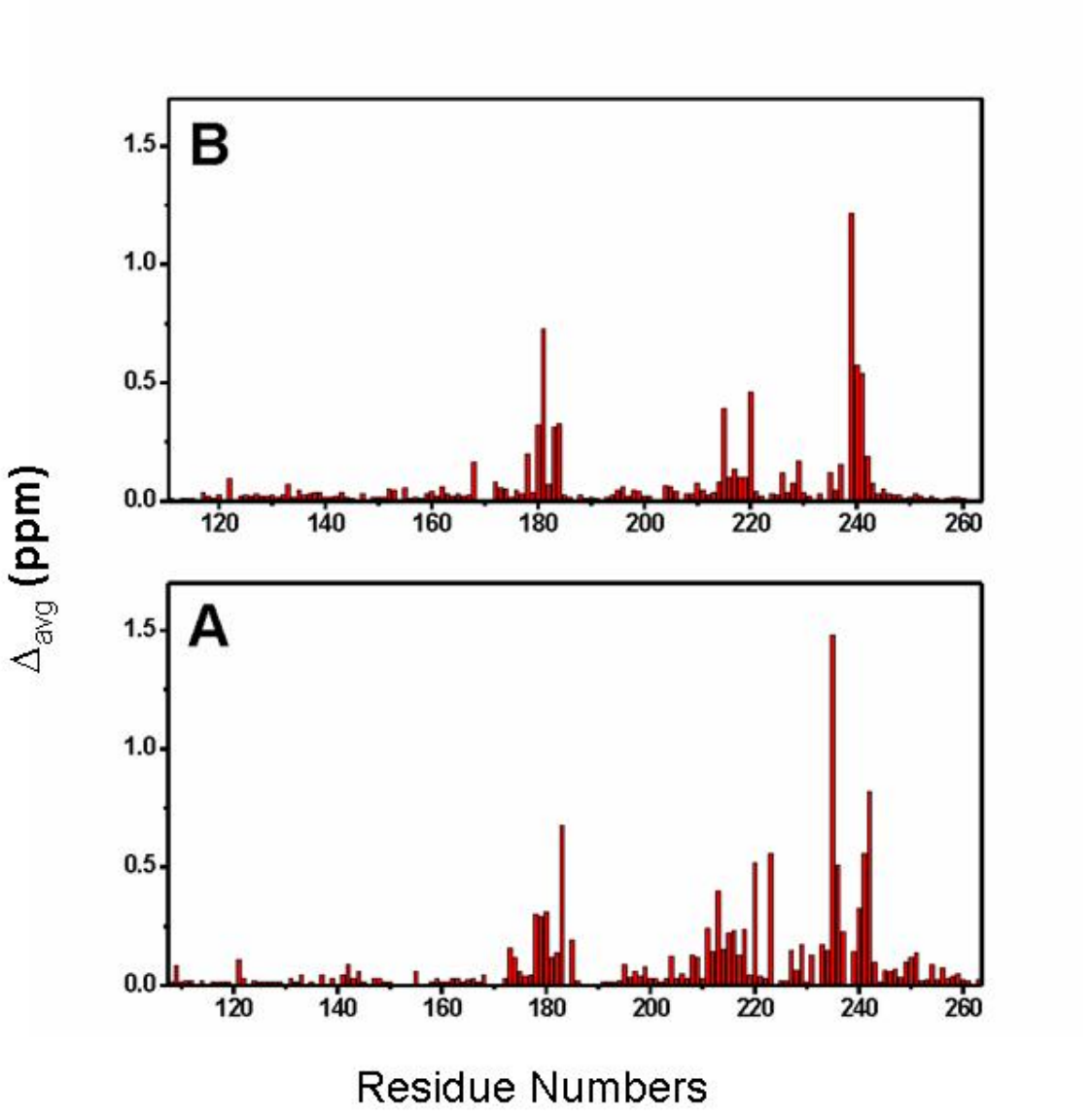


Figure 2. Garrett Plot showing effect of binding of (A) Phospinic Inhibitor and (B) NNGH to MMP12

As it is evident from Figure 2 (A) and (B), the aminoacids with significant ^{15}N-^{1}H HSQC spectral perturbations involves not only those forming the catalytic pocket but several in regions far from the catalytic site. The comparison with the Garret values obtained for the MMP-12 inhibited by NNGH, show as the both NNGH and the Phosphinic inhibitors affect the same protein regions. At the same time, the inspection of the histograms highlights a larger effect of the phosphinic inhibitor on the protein resonances.

The residues whose resonances are most affected upon binding are color-coded in Figure 3, where the solvent-accessible surface area of the protein is represented. The colors range from black ((ave > 0.5 ppm), to dark red (0.5 > (ave > 0.2 ppm), to orange (0.2 > (ave > 0.1 ppm), to yellow (0.1 > (ave > 0.05 ppm) up to white ((ave 0.05 ppm)

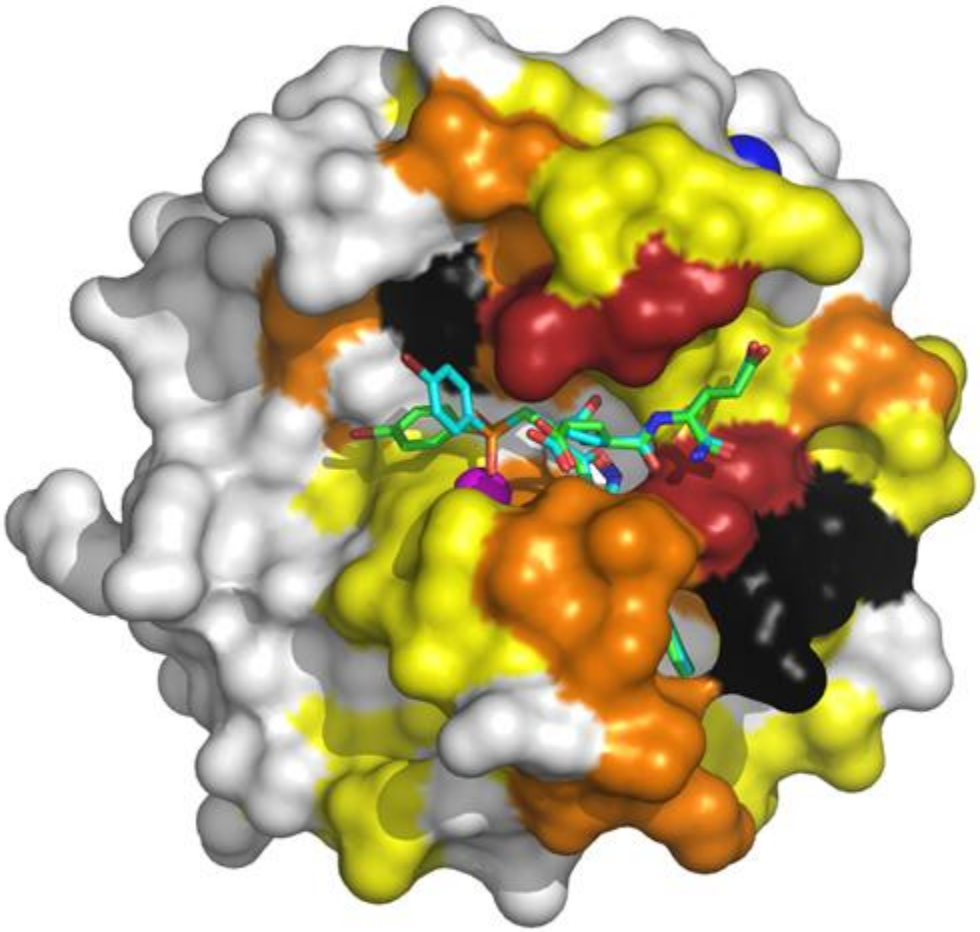


Figure 3. Solvent accessible surface area of MMP-12 complexed with the phosphinic inhibitor (X-ray structure of the socked ligand [with the right chirality!]), color-coded according to the magnitude of the chemical shift perturbation upon ligand binding.

Thermodynamics of ligand binding:

The free energy of the protein-ligand interaction is composed by enthalpic and entropic terms [(24)]. However, the analysis of such interactions on the crystal structures may account only for the enthalpic contribution. Reliable information on both, enthalpic and entropic contribution to the ligand binding can be obtained by performing Isothermal titration calorimetry (ITC) experiments [(25, 26)].

The analysis provided an accurate measurement of the ΔH^0 (+11.8 kcal/mol) associated with the binding of the phosphinic inhibitor to the protein. The endothermic peaks (Figure 4) associated with the addition of the phosphinic inhibitor show that, in this buffer, the interaction is strongly entropically driven. Indeed, the unfavorable enthalpy term must be offset by a large entropic contribution, given that the K_D is in the picomolar range (if the 0.19 nM from Dive et al., [27] is taken, a $-T\Delta S^0$ = -25 kcal/mol is obtained). This extremely high affinity makes ITC data unsuitable to estimate the binding constant of the inhibitor for MMP-12. Therefore, estimates of $-T\Delta S^0$ values were obtained by using the corrected ΔH^0 values from calorimetry and the ΔG^0 values from fluorimetry measurements.

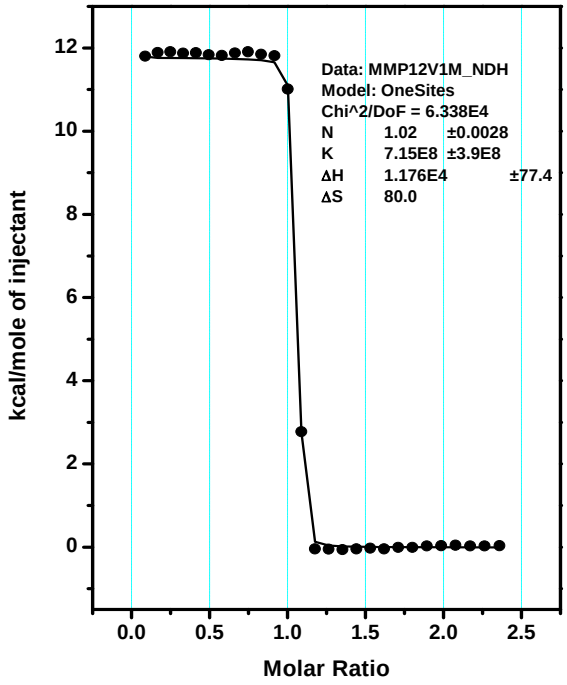


Figure 4. ITC data on the binding of the phosphinic inhibitor to MMP-12. [note that the free energy and therefore the entropy calculated from the data are meaningless, given the subnanomolar dissociation constant measured from fluorimetry].

The here presented data are consistent with a model where, upon the binding, the polar groups of the protein and of the ligand undergoes to desolvation without establishing new strong interactions. At the same time, the release of water molecules provides the large and positive entropy contribution. The extent of the entropic contribution suggests also a

relatively small loss of conformational degrees of freedom of both the drug molecule and the protein molecule upon the binding [28, 29, 30]. This hypothesis is in agreement with the analysis of the crystal structures where the conformational heterogeneity in well defined protein regions is preserved upon the binding with phosphinic inhibitor.

CONCLUSION:

The interaction of the highly selective picomolar inhibitor of MMP-12 with its protein target has been investigated in solution and in solid state. The ligand establishes several interactions with the aminoacids forming the active site and with the catalytic zinc ion. The long three-cyclic moiety fits the S1' pocket so removing the several waters molecules buried inside the cavity. The role of the two aminoacids supposed to be responsible for the observed selectivity has been demonstrated to be uninfluential in ligand binding and in selectivity. The thermodynamic investigation performed by ITC shows that the interaction is entropically driven. An extensive desolvation and a small loss of conformational entropy of both the protein and the ligand are probably responsible for the tight binding. The ability of the ligand to perform an extensive binding without drastically reducing the protein flexibility could be the origin of the selectivity for MMP-12.

REFERENCES

1. Bode, W. (2003) *Biochem. Soc. Symp.* 1–14.

2. Brinckerhoff, C. E., and Matrisian, L. M. (2002) *Nat. Rev. Mol. Cell. Biol.* 3, 207–214.

3. Egeblad, M., and Werb, Z. (2002) *Nat. Rev. Cancer* 2, 161–174.

4. Coussens, L. M., Fingleton, B., and Matrisian, L. M. (2002) *Science* 295, 2387–2392.

5. Bertini, I., Calderone, V., Cosenza, M., Fragai, M., Lee, Y. M., Luchinat, C., Mangani, S., Terni, B., and Turano, P. (2005) *Proc. Natl. Acad. Sci. U. S. A.* 102, 5334–5339.

6. Ikejiri, M., Bernardo, M. M., Meroueh, S. O., Brown, S., Chang, M., Fridman, R., and Mobashery, S. (2005) *J. Org. Chem.* 70, 5709–5712.

7. Dive, V., Georgiadis, D., Matziari, M., Makaritis, A., Beau, F., Cuniasse, P., Yiotakis, (2004) *Cell Mol. Life Sci.* 61, 2010–2019.

8. Makaritis, A., Georgiadis, D., Dive, V., and Yiotakis, A. (2003) *Chemistry* 9, 2079–2094.

9. Canales-Mayordomo, A.; Fayos, R.; Angulo, J.; Ojeda, R.; Martin-Pastor, M.; Nieto, P. M.; Martin-Lomas, M.; Lozano, R.; Gimenez-Gallego, G.; Jimenez-Barbero, J. (2006) *J.Biomol.NMR* 35, 225-239.

10. Vassiliou, S., Mucha, A., Cuniasse, P., Georgiadis, D., Lucet-Levannier, K., Beau, F., Kannan, R., Murphy, G., Knauper, V., Rio, M. C., Basset, P., Yiotakis, A., and Dive, V. (1999) *J. Med. Chem.* 42, 2610–2620.

11. Keller, R. L. J. The Computer Aided Resonance Assignment Tutorial, 1.3; Cantina: Verlag, 2004.

12. Leslie, A. G. W. (1991) in *Molecular data processing* (Moras, D., Podjarny, A. D., and Thierry, J.-C., Eds.) pp 50-61, Oxford University Press, Oxford. S13

13. Evans, P. R. (1997) SCALA, *Joint CCP4 and ESF-EACBM Newsletter* 33, 22-24.

14. Tickle, I. J. and Driessen, H. P. C. (1996) *Methods Mol. Biol.*, 56, 173-203.

15. Rossmann, M. G., and Blow, D. M. (1962) *Acta Cryst.* 15, 24.

16. Crowther, R. A. (1962) *The Molecular Replacement Method* (Rossmann, M. G., ed.) Gordon & Breach: New York.

17. Vagin, A.; Teplyakov, A. (1997) *J.Appl.Crystallogr.* 30, 1022-1025.

18. Vagin, A.; Teplyakov, A. (2000) *Acta Crystallogr.D Biol.Crystallogr.* 56, 1622-1624.

19. Murshudov, G. N.; Vagin, A. A., Dodson, E. J. (1997) *Acta Cryst. D* 53, 240-255.

20. McRee, D. E. (1992) *J.Mol.Graphics* 10, 44-47.

21. Lamzin, V. S. (1993) *Acta Crystallogr.D Biol.Crystallogr.* 49, 129-147.

22. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283-291.

23. Indyk, L., Fisher, H.F. (1998) *Energetics of Biological Macromolecules, Pt B* 295 350-364.

24. Andrews, P. R., Craik, D. J., Martin, J. L. (1984) *J Med Chem* **27,** 1648-1657.

25. Baker, B. M., Murphy, K. P. (1997) *J Mol Biol* **268,** 557-569.

26. Appella, E., Anderson C.W. (2005) *FEBS J.* **272**, 5099-100.

27. Laurent, D., Vassilis R., Arnaud D., Anastasios M., Fabrice B., Philippe C., Athanasios Y., Vincent D. (2006) *J. Biol. Chem.* 281 (16), 11152-11160.

28. Bertini, I., Calderone, V., Fragai, M., Giachetti, A., Loconte, M., Luchinat, C., Maletta, M., Nativi, C., Yeo, K. J. (2007) *J Am. Chem Soc.* **129,** 2466-2475.

29. Breslauer, K. J., Remeta, D. P., Chou, W. Y., Ferrante, R., Curry, J., Zaunczkowski, D., Snyder, J. G., Marky, L. A. (1987) *Proc Natl Acad Sci U S A* **84,** 8922-8926.

30. Chaires, J. B. (1997) *Biopolymers* **44,** 201-215.

7. General Discussion & Future Perspectives

General Discussion:

The final aim of PhD thesis work can be described even with the sentence "Understanding the molecular mechanistic of the Inhibitors/Substrates-MMP12 interaction and to improve the drug designing strategies for this protease".

Cat, Hpx and Full-length MMP12 (wild type and various mutants) were expressed during the course of PhD. The protocols and strategies developed during the project resulted in better yield of purified stable samples of protein. These samples allowed us to deeply investigate the molecular mechanism of MMP12. Catalytic activation as well the interaction. If it is true that we focused our study on MMP12, then our findings might be applicable to whole MMP family or any structurally characterized target. We give an important contribution to the understanding of the mechanicistic features of MMP12 as well as MMPs. The present studies are a significant example of the synergy between various biophysical and biochemical techniques.

The relatively mobility of Cat and Hpx like domains has been verified which led to the possibility that domain flexibility can mediate the activation of the enzyme. These results can be representative for other members of MMP family having the short linker region between the two domains and help to understand the role of mobility on substrate recognition and on catalytic mechanism. Moreover the preliminary results of the interaction studies of MMP12 full length as well Cat and Hpx domain individually with physiological substrate elastin and its subunits are exciting and fruitful in understanding the long debated role of Hpx domain in binding and catalysis of substrate for this enzyme. This provides us the possibility to deeply investigate the mechanism of substrate recognition by MMP12 and in general similar class of MMPs.

The knowledge of the catalytic mechanism of MMPs is a key-tool for understand how to design more selective inhibitor. The presented data by combining the structural information with thermodynamic studies provided hints for the possibility of designing selective inhibitor with increased affinity for MMP12. This structural determinant studies helped to have a better looks insight of the interaction and thus inhibition. Further the structure based drug designing led to development of biotinylated-probe for profiling

MMP12 in particular and some other member of the MMP family. Such activity based bifunctional probes could be of high importance and efficacy *in vivo* and *in vitro*.

Future Perspective:

The initial results for interaction studies of MMP12 and Elastin and its two components are interesting and exciting. The future course of work is to understand the molecular mechanistic and to address the potential role of Hpx role in substrate recognition by MMP12 and the structural determinants responsible for the interactions. The long term perspective is to investigate the mechanism of activation of full length MMP12 (Pro-Cat-Hpx) and to verify the inter domain mobility for other members of MMP family.

Further, activity based probes for profiling of various MMPs in different pathological conditions and designing of selective inhibitors for metalloelastase as well as other members of family is one of the important parts of future work.