University of Florence

International Doctorate in Structural Biology

Cycle XIX (2004-2006)

Development of expression and refolding methods of human matrix metalloproteinases for drug discovery application

Ph.D. thesis of

Kwon Joo Yeo

Tutor

Prof. Ivano Bertini

Coordinator

Prof. Claudio Luchinat

S.S.D. CHIM/03

Contents

1 Introduction

1.1 Matrix metalloproteinases (MMPs)

The extracellular matrix (ECM) plays a critical physiological role for many cellular function of the organism [1-2]. Many of the extracellular signaling events that regulate cell behaviour occur at or near the cell membrane and are regulated by proteolytic processes [3].

MMPs are a family of Zn^{2+} and Ca^{2+}-dependent endopeptidases that play a pivotal role in the degradation and remodeling of ECM. Collectively, they are able to digest all components of ECM and to regulate the cell-matrix interactions (Figure 1). The ability of the activated MMPs to degrade ECM and to cleave cell surface molecules influences diverse physiologic processes such as wound repair, apoptosis and tissue remodeling that take place during embryonic development and reproduction [4]. Under loss of their regular functions, MMPs are thought to promote tumor expansion and invasion so favoring the movement of cancer cells across the vascular membrane, the local growth and invasion of secondary tumors, and the formation of new blood vessels, which is an important requisite for tumor growth (Figure 2) [5-11]. MMPs also are implicated in many other diseases such as arthritis, atherosclerosis, tissue ulcers and fibrosis [12-14]. Due to their role in cancer progression, several MMPs are considered important drug targets for anticancer therapy. Over the past decade, significant advances have been made in the field of MMP research. These included a better understanding of the biochemistry of these enzymes in terms of their activation, regulation and substrate specificity, the determination of their structure by X-ray crystallography and NMR, and the design of orally-available inhibitors [15-19]. Unfortunately the clinical trials of the inhibitors designed so far have provided disappointing results. In fact in spite of the encouraging preclinical efficacy in pharmacological models, side-effects have been observed in the patients treated with these compounds. The lack of selectivity is considered at the origin of the observed side-effects and of the poor efficacy [8,

18]. Therefore, the major challenges in MMP research are a better understanding of the complex role of these enzymes in human diseases, and the design of selective inhibitors for the clinical use.

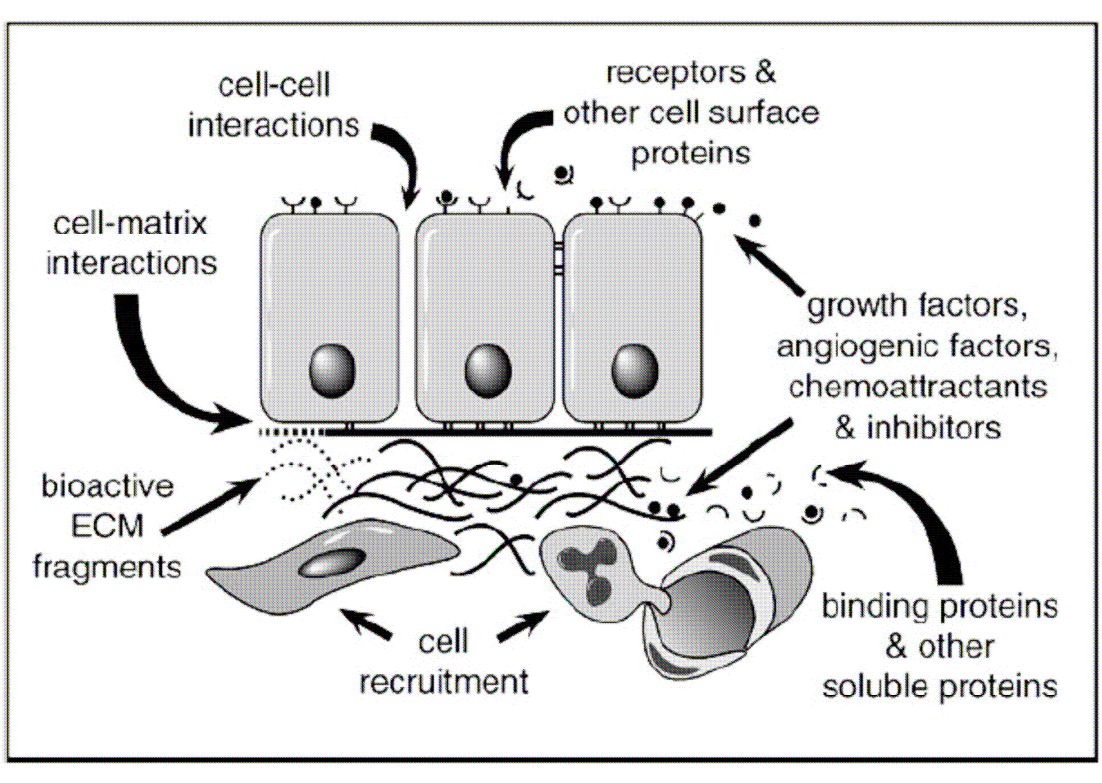


Figure 1. Potential mechanisms of MMP-mediated cellular signaling (reprinted from ref. 17)

1.2 Structural features of MMPs

To date, 24 different vertebrate MMPs have been identified. Of these, 23 have been found also in humans [20-22]. A typical full length MMP consists of a signal peptide, a prodomain, a catalytic domain, a linker peptide (called ‘hinge region’) and a C-terminal hemopexin domain (Figure 3 and 4). A catalytic domain and a C-terminal hemopexin domain are connected by a linker peptide. The hemopexin domain is absent in MMP-7, MMP-26 and MMP-23. MMP-23 has a unique cystein-rich domain and an immunoglobulin-like domain. Gelatinase A (MMP-2) and gelatinase B (MMP-9) have three fibronectin

type II motifs inserted in the catalytic domain. MMPs are produced in their inactive zymogenic forms (pro-MMPs), which are subsequently proteolytically activated in an elaborate set of events. The latent MMPs (pro-MMPs) contain a conserved cysteine residue called “cysteine switch”, that interacts with the catalytic zinc ion, and maintains the enzymes in latent state. MMPs are activated *in vivo* in the extracellular space by the removal of the prodomain through the action by other MMPs, or *in vitro* by organomercurial compounds or different proteinases such as trypsin, furin and plasmin [23-25].

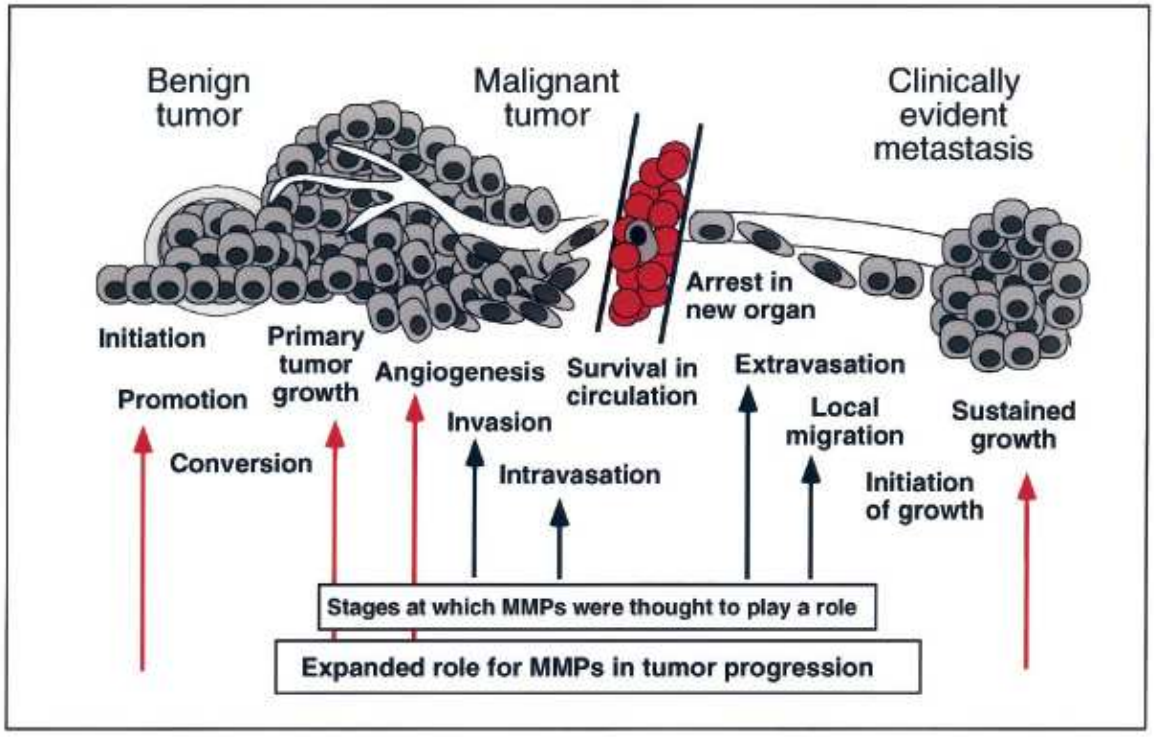


Figure 2. Matrix metalloproteinases (MMPs) in tumor progression. The cellular changes that occur as a normal cell becomes a benign, malignant, and metastatic tumor are depicted. MMPs have been classically thought to contribute to the tissue destruction required for cells to invade, intravasate, extravasate, and migrate. More recent evidence suggests these enzymes can also play a role in the growth of benign and malignant tumors, angiogenesis, and the sustained growth of metastatic lesions (reprinted from Nelson et al, *J. Clin. Oncol.*, 2000, 18, 1135-1149

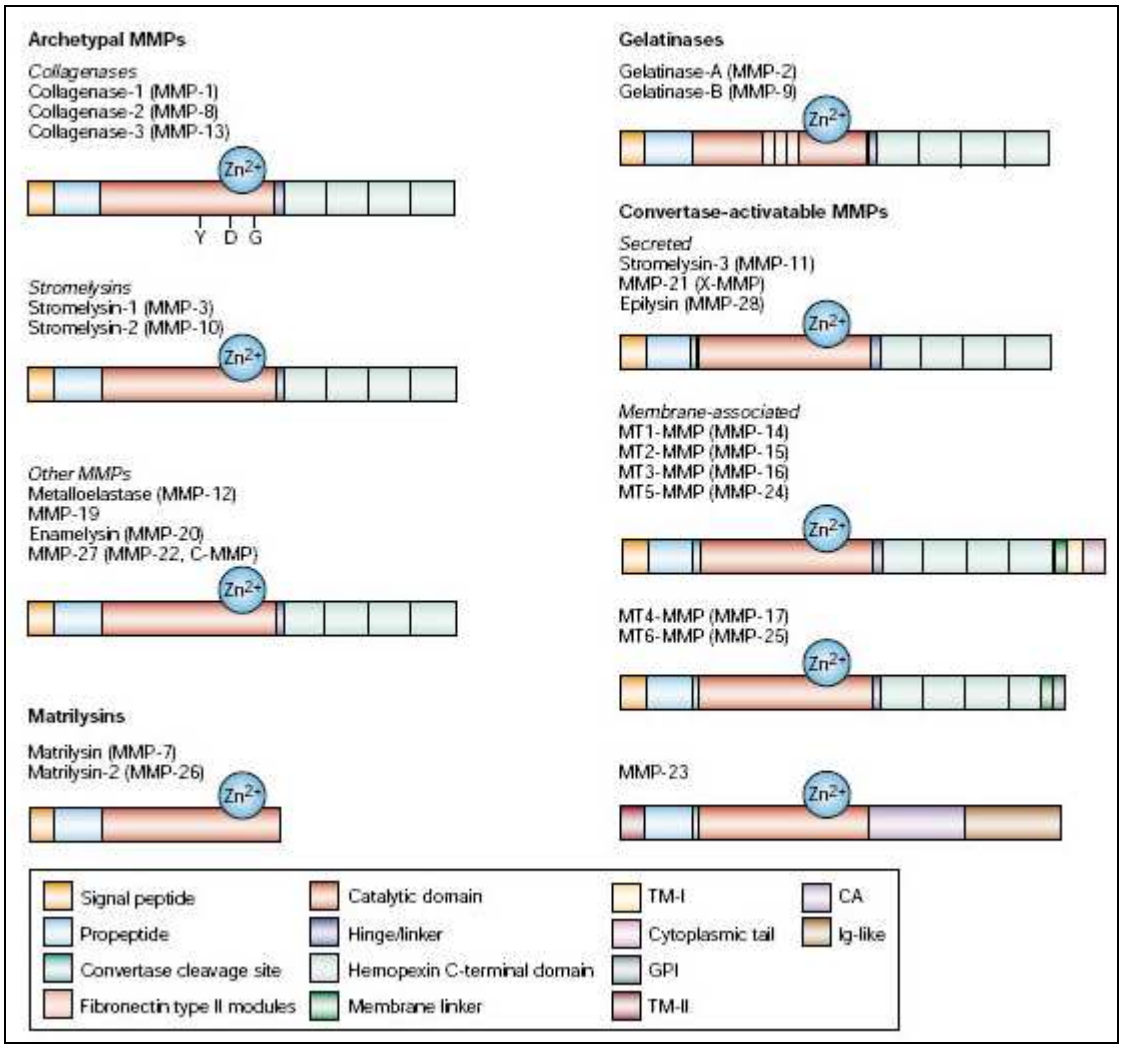


Figure 3. Schematic representation of the structure of the 23 human matrix metalloproteinases (MMPs), which are classified into different groups on the basis of domain organization. Archetypal MMPs contain a signal peptide (necessary for secretion), propeptide, a catalytic domain that binds zinc (Zn^{2+}) and a hemopexin carboxy (C)-terminal domain. Y, D, and G represent tyrosine, aspartic acid and glycine amino acids that are present in the catalytic domain of all collagenases. Matrilysins contain the minimal domain organization that is required for secretion, latency and catalytic activity. Gelatinases contain fibronectin type II modules that improve collagen and gelatin degradation efficiency. Convertase-activatable MMPs contain a basic insert in the propeptide that is targeted by furin-like proteases (convertase cleavage site). MMPs that belong to this group can be secreted enzymes, or membrane-anchored via GPI (glycosylphosphatidylinositol), type I or type II transmembrane (TM) segments. MMP-23A and MMP-23B contain unique cysteine array (CA) and immunoglobulin (Ig)-like domains in their C-terminal region (reprinted from Overall et al, *Nat. Rev. Cancer*, 2002, 2, 657-672).

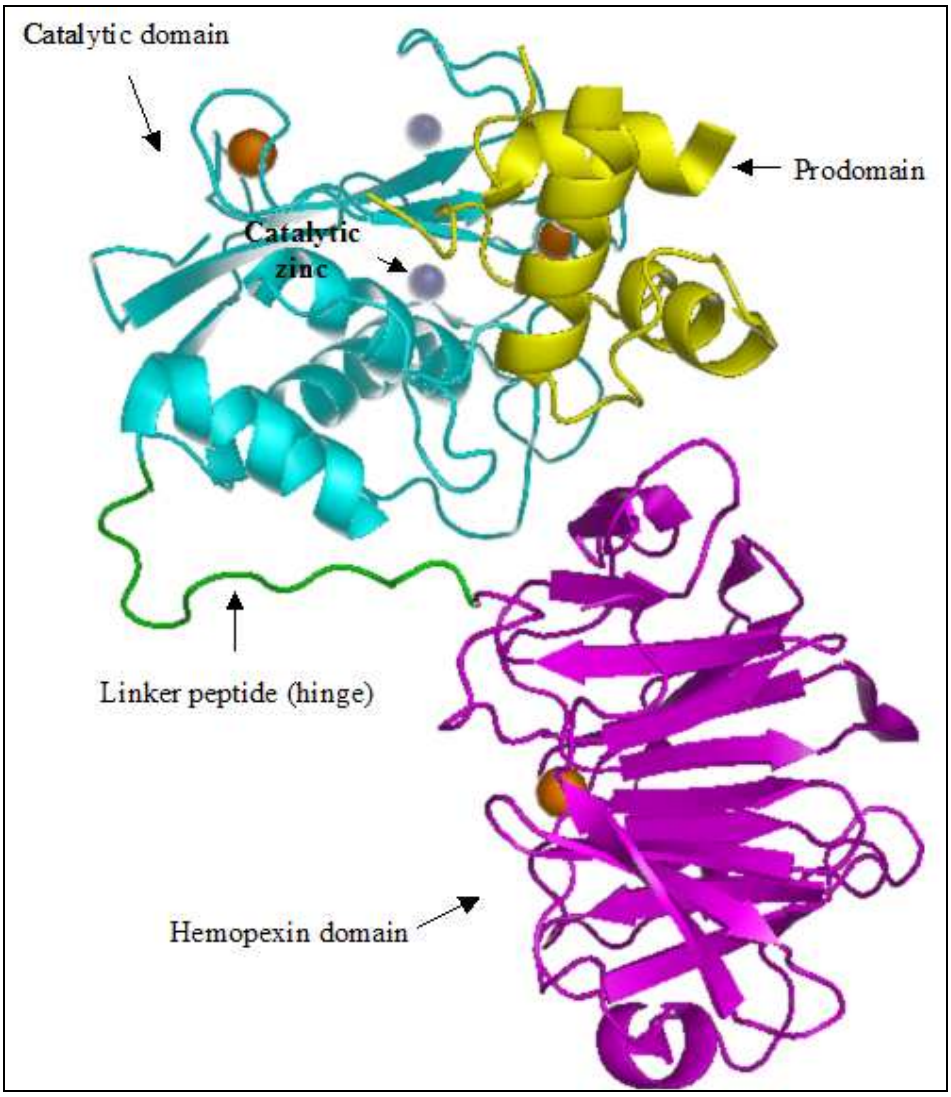


Figure 4. The crystal structure of the full length of human MMP-1 (PDB code: 1SU3)

The catalytic domain contains an highly conserved zinc-binding sequence HEXXHXXGXXH, which is essential for the proteolytic activity of MMPs. Glutamate and aspartic acid rich sequences at the N- and C-terminal ends of the catalytic domain are thought to represent calcium-binding motifs. The catalytic domain also contains a conserved methionine, forming a "Met-turn" after the zinc- binding motif, which forms a base to support the structure around the catalytic zinc. With the exception of MMP-7, MMP-26 and MMP-23, all MMPs have an hemopexin domain that shows sequence similarity to the heme-

binding protein hemopexin, and it is highly conserved among MMPs. Two cysteine residues in the hemopexin domain form a disulfide bridge folding the domain into a four bladed propeller structure [26-27]. The hemopexin domain plays a role in regulating both substrate specificity and proteolytic activities. It is also important for the binding of tissue inhibitors of metalloproteinases (TIMPs) [26].

1.3 Members of the MMPs and their biological functions

On the basis of substrate specificity, domain arrangement and sequence similarity, MMPs are grouped into collagenases, gelatinases, stromelysins, matrilysin, membrane-type (MT)-MMPs and others (Table 1).

The human collagenases are MMP-1, MMP-8 and MMP-13. These MMPs are capable of degrading triple helical fibrillar collagens type I, II and III and of digesting other matix and non-matrix proteins [21, 28]. Also MMP-2 and MT1-MMP (MMP-14) can cleave fibrillar collagens [29-30]. Recent studies show that collagenases bind and partially unwind triple helical structure of native collagens before the hydrolysis of the peptide bonds, so suggesting that the hemopexin domain is essential for the catalytic mechanism [30-31]. Therefore, the catalytic domain and the hemopexin domain, have an important role in collagen degradation. Type I collagen is cleaved by several MMPs at specific sites and more exactly between Gly775-Ile776 of the α1(I) chain and Gly775-Leu776 of the α2(I) chain. These selective cleavages produce 3/4 N-terminal and 1/4 C-terminal collagen fragments, which can be denatured spontaneously into gelatin at body temperature, and can be further degraded by other MMPs [32]. MMP-1 can also activate the exodomain of protease activated receptor 1 (PAR-1) by cleaving the Arg-Ser bond [10, 33].

Gelatinases (MMP-2 and MMP-9) readily digest the denatured collagen and gelatin. They also digest a number of ECM components including type IV, V and XI collagens, laminin, aggrecan core proteins, etc. [34-35]. MMP-2, but not MMP-9, has collagenolytic

Table 1. Human matrix metalloproteinases and their substrates

Enzyme	Common name(s)	Major substrates
MMP-1	Collagenase-1, interstitial collagenase	Collagen I, II, III, X, gelatin
MMP-2	Gelatinase A,	Gelatin, laminin, collagen I, IV, V, VII, fibronectin, elastin
MMP-3	Stromelysin-1	Proteoglycans, fibronectin, gelatin, collagen III, IV, V, IX
MMP-7	Matrilysin-1	Gelatin, fibronectin
MMP-8	Collagenase-2, neutrophil collagenase	Collagen I, II, III, X, gelatin
MMP-9	Gelatinase B	Gelatin, proteoglycans, collagen IV, V, VII,
MMP-10	Stromelysin-2	fibronectin, elastin
MMP-11	Stromelysin-3	Gelatin
MMP-12	Macrophage Metalloelastase	Gelatin, Elastin
MMP-13	Collagenase-3	Collagen I, II, III
MMP-14	MT1-MMP	Collagen, aggrecan
MMP-15	MT2-MMP	Collagen, aggrecan
MMP-16	MT3-MMP	
MMP-17	MT4-MMP	
MMP-19	RASI-1	Gelatin, fibronectin
MMP-20	Enamelysin	
MMP-21	XMMP	
MMP-23	CA-MMP	
MMP-24	MT5-MMP	
MMP-25	MT6-MMP	
MMP-26	Matrilysin-2, Endometase	
MMP-27		
MMP-28	CMMP, Epilysin	

activity [30]. MMP-2 is activated on the cell surface through a unique multi-step pathway that involves MMP-14 (MT1-MMP) and the tissue inhibitor of metalloproteinase 2 (TIMP-2) [36-37]. MMP-9 has been shown to mobilize vascular endothelial growth factor and to promote angiogenesis. In the K14-HPV16 carcinogenesis model, MMP-9 expression by inflammatory cells, leads to a keratinocyte hyperproliferation, angiogenesis and to an

increased incidence of squamous cell carcinomas [6].

The stromelysins (MMP-3, MMP-10 and MMP-11) display a broad ability in degrading ECM proteins. Conversely they are unable to cleave the triple helical fibrillar collagens. MMP-3 and MMP-10 are similar in structure and in substrate specificity and they participate to proMMP activation [38].

In matrilysins (MMP-7 and MMP-26) the hemopexin domain is absent. MMP-7 is synthesized by epithelial cells and is secreted apically. Besides ECM components MMP-7 processes cell surface molecules such as pro-α-defensin, Fas-ligand, pro-tumor necrosis factor α and E-cadherin. MMP-26 is expressed in normal cells such as those of the endometrium and in some carcinomas. It digests several ECM molecules, and unlike most other MMPs, it is largely stored intracellularly [39].

MT-MMPs, Type I transmembrane proteins (MMP-14, MMP-15, MMP-16 and MMP-24) and glycosylphophatidyinositol-anchored proteins (MMP-17 and MMP-25), have a furin recognition sequence at the C-terminal of the propeptide (Figure 3). The importance of these enzymes is related to their ability of activating pro-MMP-2 on the cell surface. With the exception of MT4-MMP (MMP-17), all MT-MMPs can activate proMMP-2. MT1-MMP serves as cell membrane receptor for the complex formed between the latent form of MMP-2 (pro-MMP-2) and the tissue inhibitor of matrix metalloproteinases 2 (TIMP 2). In this way pro-MMP-2 is brought on the cell surface where another MT-MMP, free of TIMP, can cleave out the prodomain [17, 36]. MT1-MMP (MMP-14) has also collagenolytic activity on collagens type I, II and III [40].

The other seven MMPs are not classified into the above categories. MMP-12, called macrophage metalloelastase, is mainly expressed by alveolar macrophages [41] and digests elastin, other ECM components [42] and non-matrix proteins such as myelin basic protein and α1-antitrypsin. It is also able to process tumour necrosis factor-a (TNF) to its mature form [43]. It is essential also for macrophage migration [44]. MMP-19 digests many

ECM molecules including the components of basement membranes [45] and it is also called RASI (rheumatoid arthritis synovial inflammation) as it is found in the activated lymphocytes and in the plasma of patients with rheumatoid arthritis. MMP-19 is also recognized as an autoantigen in patients with rheumatoid arthritis and systemic lupus erythematosis [46]. MMP-20, called also Enamelysin, is primarily located within newly formed tooth enamel and it is able to digest amelogenin [47] MMP-21 was originally found in Xenopus [48] and more recently in mice and human [49]. It is expressed in various fetal and adult tissues and in basal and squamous cell carcinomas [50]. It digests gelatin, but information about other activities towards the ECM components are not available. MMP-23 is a unique member as it has unique C-terminal cysteine-rich immunoglobulin-like domains instead of the hemopexin domain and lacks the cysteine-switch motif of the prodomain [51]. MMP-27 was first found in chicken embryo fibroblasts [52]. Chicken MMP-27 digests gelatin and casein and causes autolysis of the enzyme, but little information are available on the activity of mammalian enzyme. The last MMP to be identified has been the Epilysin or MMP28, mainly expressed in keratinocytes. Expression patterns in intact and damaged skin suggest that this MMP might function in tissue homeostasis and wound repair [53].

1.4 MMPs as drug targets

In general, MMPs are secreted as zymogens that are activated in vivo by proteinases including MMPs [21]. Low pH and heat treatment can also lead to activation of pro-MMPs. Since MMPs were activated, the proteolytic activity of mature enzymes is controlled by physiological inhibitors. As there is growing experimental evidence that malignant tumors utilize MMPs for tumor growth and spreading, inhibitors of MMPs may represent a new class of anticancer drugs [18, 54-55]. Native inhibitors of MMP activity are serum proteins such as α-2 macroglobulin and, more specifically, the family of tissue inhibitors of metalloproteinases (TIMPs). To date, four members of the family, TIMP-1,

TIMP-2, TIMP-3, and TIMP-4 have been characterized. These proteins are able to inhibit the activity of all known matrix metalloproteinases [36, 56-57]. However, the therapeutic use of these proteins is likely to be limited because of their low oral bioavailability and because of the limited tissue penetration. Synthetic inhibitors ideally have the desirable properties of high water solubility, oral bioavailability, and low toxicity. So far, the most extensively used inhibitor in clinical trials is marimastat, a hydroxamic acid derivative which exhibits broad-spectrum inhibitory activity against MMPs [54]. However, the use of broad-spectrum MMP inhibitors in clinical trials has been disappointing, primarily because of lack of efficacy and survival benefits [8]. For these reasons, it is necessary to achieve a selective inhibition of each human MMP. Recently, to design selective inhibitors, the attention has been focused not only towards the S_1' pocket of catalytic site but also towards the loop regions (Figure 5) [18, 58-60]. The S_1' pocket plays an important role in determining P_1' specificity but the mobility of loop regions is also important since it allows the protein to accommodate a variety of inhibitor moieties. Therefore, in the future, an accurate comparison of the S1' pockets and a deep investigation of the loop regions will provide valuable data for a rational design of specific inhibitors. In this context the availability of high resolution structures is of great help.

More recently, it is suggested that the investigation of the structural features of the hemopexin domain and of the linker peptide is important to understand the physiological role and for drug design [55,61]. Up to now, the functional roles of the hemopexin domain and that of the linker peptide of MMPs have not been completely clarified. In the case of collagenases (MMP-1, MMP-8, MMP-13), the hemopexin domain seems to mediate binding of the enzyme to the native collagen [62-64]. Therefore, the hemopexin domain and linker peptide of MMPs have an important role in regulating the degradation of their substrates (ECM components). Indeed, from the point of view of drug design, targeting sites contributing to the interaction with the substrates but far from the active site could be a good

strategy to enhance the selectivity of candidate drugs.

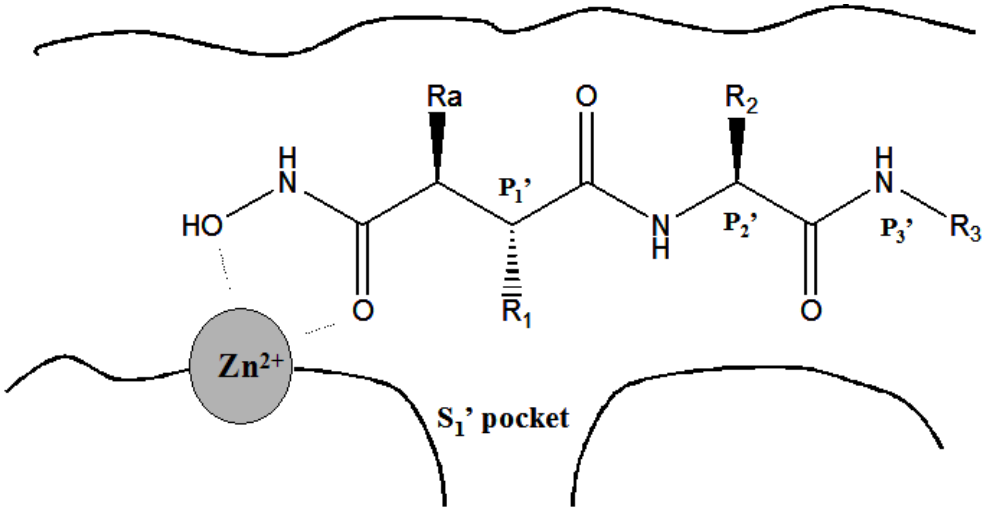


Figure 5. Schematic representation of the active site cleft of MMPs complexed with hydroxamic acid inhibitor. Ra, α substituent; R_1, P_1' substituent; R2, P2' substituent; R3, P3' substituent.

1.5 Aim of the research

To accomplish the rational drug design, high resolution structures of the proteins are highly required. To analyze the structural details of each MMP, and the subtleties of the ligand-protein interactions, the availability of highly concentrated, soluble and stable protein samples is needed. Moreover, to investigate the interaction of the matrix metalloproteinases with physiological substrates (e.g. models of extracellular receptors or extracellular matrix components) the availability of full-length proteins is of fundamental importance. The present PhD project has been focused on the development of new methods for the expression, purification and refolding of recombinant human MMPs. Several catalytic domains of

human MMPs (MMP-1, MMP-7, MMP-8 and MMP-12), one full length active protein (MMP-12) and the hemopexin domain of MMP-12 have been obtained so far. Inactive mutants of the catalytic domain of MMP-1, MMP-12, and the full length active construct of MMP-12 (included an inactive Cd^{2+}-substituted form of MMP-12 full length) have been also produced. For the full length and the hemopexin domain of MMP-12, indeed, it is hard to produce a large amount of the proteins because aggregation and precipitation occur during the refolding steps. Fortunately, large amounts of the proteins with well-folded and soluble state have been produced using the new methods developed during the present PhD research. Thanks to these methods, the solution structure of the hemopexin domain of MMP-12 has almost been solved and the solution structure of the full length of MMP-12 is in progress. Therefore, we expect that the solution structures of active full length MMP-12 will provide important clues for interaction studies of MMP with substrates such as collagen, elastin and for the design of selective inhibitors.

2 Methods

2.1 Expression, purification and refolding of human MMPs

The development of new methods for MMP cloning, expression and purification that were the goals of this PhD project has been described in details in Chapter 3.
All the proteins expressed during these research activities have been produced according the general strategy described below.

The cDNAs encoding the sequences for the different proteins were amplified from human cDNAs by polymerase chain reaction (PCR). The selected constructs were cloned into the pET21 expression vector (Novagen) using *Nde I* and *Xho I* or *Nde I* and *Hind III* as restriction enzymes (BioLabs). The recombinant vectors were transformed into *Escherichia coli* strains, BL21(DE3) Codon plus or in BL21(DE3). Single point mutations were performed using the QuickChange™ Site-Directed Mutagenesis Kit from Stratagene. The pET21 has been selected as plasmid vector since its efficiency in the expression of the catalytic domain of MMPs. All the proteins were expressed at 37 ^{0}C using 0.5 mM of IPTG to induce the gene transcription. The expression was carried out in rich media (2 x YT) or in minimal media for ^{15}N- and ^{13}C-labeled proteins. The proteins, precipitated in the inclusion bodies, were solubilized after lysis of the cells in a solution containing 8 M urea or 6 M guanidine hydrochloride. The protein was purified on the ion exchange columns, using a linear gradient of NaCl up to 0.35 M. The purified proteins were then refolded by using a multi-step dialysis against a buffer containing Tris, $CaCl_2$, $ZnCl_2$, NaCl, and acetohydroxamic acid (AHA).

2.2 Crystallization

The crystals of the proteins (MMP-8, the active form of MMP-12, and the inactive form of MMP-12) were obtained by the hanging drop vapor diffusion method. Crystals of

MMP-12 were obtained as previously reported [66], while those of MMP-8 were grown at 20 ^{0}C from a solution containing 0.1 M Tris-HCl, 20% PEG-3350, 0.2 M $MgCl_2$ at pH 8. In both cases the crystallization buffer contained 200 mM of the weak inhibitor acetohydroxamic acid (AHA). To obtain the active uninhibited enzymes, MMP crystals were then extensively washed with the same crystallization buffers lacking AHA. The ProGlnGlyIleAlaGly peptide (INBIOS s.r.l., Naples) was soaked into the crystals for 1-3 days in order to obtain the two- or the one-peptide adducts.

2.3 Enzyme assays

Fluorimetric assay was performed in 50 mM HEPES buffer (pH 7.0), with 10 mM $CaCl_2$, 0.05% Brij-35 and 0.1 mM $ZnCl_2$ using 1 nM of enzyme and 1 μM of the fluorescent-quenched peptide substrate Mca-Pro-Leu-Gly-Leu-Dpa-Ala-Arg-NH2 (Biomol Inc.) at 298 K. The inhibition constants for the investigated inhibitors were determined evaluating their ability to prevent the hydrolysis of the peptide. Activity tests were performed by following the UV absorption in a buffer containing 50 mM HEPES (pH 7.0), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.05 % Brij-35, 1mM DTNB (5,5'-dithiobis-2-nitrobenzoic acid) at 25 ^{0}C, 100 uM of Ac-Pro-Leu-Gly-[2-mercapto-4-methyl-penntanoyl]-Leu-Gly-OC_2H_5 as substrate and 50 nM protein.

2.4 Cadmium substitution

The sample of the full length of Cd^{2+}-MMP12 was prepared by exhaustive dialysis of the Zn^{2+}-MMP12 against a buffer containing 0.3 mM $CdCl_2$. The Zn^{2+}-MMP12 was firstly exchanged by dialysis against a buffer containing 10 mM HEPES (pH 6.7), 5 mM $CaCl_2$, 0.3 mM $CdCl_2$; 0.3 M NaCl and 0.2 M AHA. The final sample for structural studies was dialyzed against a buffer containing 20 mM Tris (pH 7.2), 10 mM $CaCl_2$, 0.1 mM $CdCl_2$, 0.1 M NaCl and 0.2 M AHA.

2.5 NMR sample preparation

The NMR samples of the catalytic domains were obtained by dialyzing the protein against a buffer containing 10 mM Tris (pH 7.2), 10 mM $CaCl_2$, 0.1 mM, $ZnCl_2$, 0.3 M NaCl, 0.2 M AHA, 10% D_2O. In case of the hemopexin domain of MMP-12, the protein was dialyzed against a buffer containing 20 mM Tris (pH 7.2), 10 mM $CaCl_2$, 0.15 M NaCl, 10% D_2O. The NMR sample of the full length of MMP-12 was prepared by dialyzing the protein against a buffer containing 10 mM Tris (pH7.2), 10 mM $CaCl_2$, 0.1 M NaCl, 0.2 M AHA, excess of N-isobutyl-N-[4-methoxyphenylsulfonyl]glycyl hydroxamic acid (NNGH), 10 % D_2O. In case of the full length Cd^{2+}-MMP12 sample, 0.1 mM $CdCl_2$ was added instead of $ZnCl_2$.

3 Results

3.1 Catalytic domains of human MMPs

MMP-1

Pro-cat MMP-1 (Pro21-Pro269) was cloned into pET21 expression vector using *NdeI* and *XhoI* as restriction enzymes and expressed in BL21(DE3) Codon Plus cells. For the ^{15}N-labeled protein, the test expression was performed in minimal media at 37 ^{0}C (Figure 6. A). Most of the proteins were expressed as inclusion bodies, even if some protein was soluble at 25 ^{0}C. After induction with 0.5mM IPTG, the cells were harvested for 4hrs. The bacterial pellets were collected by centrifugation and lysed by using the following protocol.

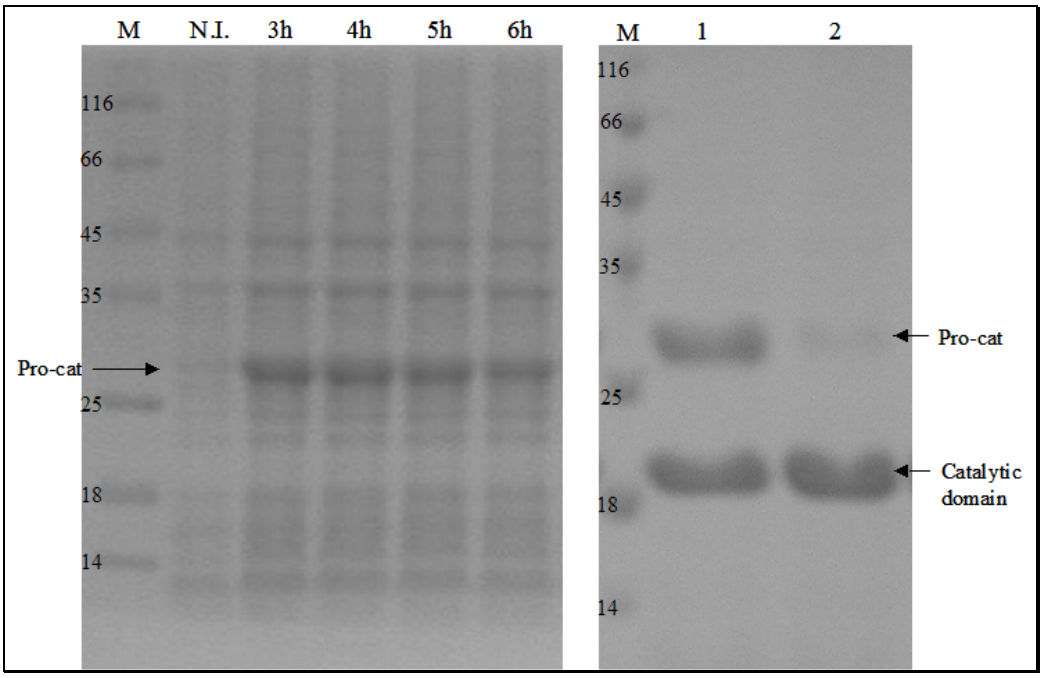


Figure 6. SDS-PAGE (sodium dodecyl sulphate polyacrylamide gel electrophoresis) analysis of test expression of the pro-cat MMP-1 (Pro21-Pro269) in minimal media at 37 ^{0}C (A) and activation of the refolded protein (B; lane 1, after refolding; lane 2, after activation by 1 mM APMA at 4^{0}C overnight).

The bacterial pellets were suspended in a lysis buffer containing 50mM Tris (pH 8.0), 25% sucrose, 0.1 M NaCl, 0,2 mM EDTA, 1mM DTT, and then stirred with some lysozyme for 20 min at 4 ^{0}C. The same volume of a buffer containing 2% Triton, 50mM Tris (pH 8.0), 0.1 M NaCl, 0,2mM EDTA, 1mM DTT was added, and then the viscous solution was sonicated several times in cold room. After sonication, the solution was stirred with 1 mg/ml DNase and 100 mM $MgCl_2$ for 20 min at 4 ^{0}C. The pellets were collected by centrifugation, and then washed three times by suspension in 50mM Tris (pH 8.0), 5mM EDTA, 1mM DTT and followed by centrifugation. The pellets containing inclusion bodies were solubilized in 2 M urea, 20 mM Tris (pH 8.0).

The protein was purified on the Hitrap Q column (Pharmacia) with a buffer containing 2 M urea and 20 mM Tris (pH 8.0). The elution was performed using a linear gradient of NaCl up to 0.35 M. The purified protein was then refolded by using a multi-step dialysis. The refolded protein was activated by 1 mM APMA (4-aminophenylmercuric acetate) at 4 °C overnight (Figure 6. B). The catalytic domain of MMP-1 was purified using size-exclusion chromatography with a buffer containing 0.2 M AHA, and then dialyzed against a proper buffer for structural or biological studies. The detailed procedures and results are described in chapter 4.1. However, pro-cat MMPs require the activation of the protein and a gel filtration to remove prodomain, Unfortunately, we couldn’t obtain any crystal from the protein derived by removing the prodomain. Therefore, a different construct of the catalytic domain of MMP-1 (Asn106-Gly261) was cloned and expressed as described below.

The catalytic domain of MMP-1 (Asn106-Gly261) was cloned and expressed using the same method used for the pro-cat MMP-1. The test expression of the protein was performed in minimal media at 37 ^{0}C. SDS-PAGE analysis of the test expression of the proteins is shown in Figure 7.

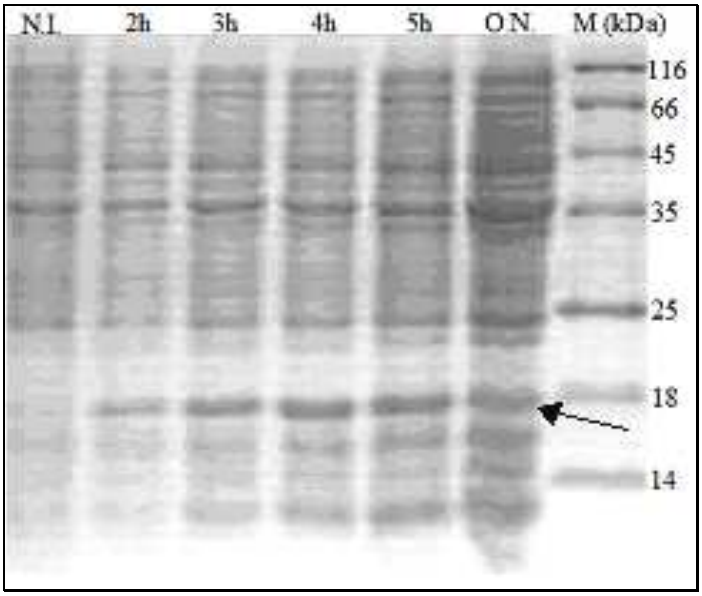


Figure 7. SDS-PAGE analysis of test expression of the catalytic domain of MMP-1 in BL21(DE3) Codon Plus cells at 37 ^{0}C

In order to maximize the production of the protein, the incubation was continued at 37 ^{0}C for 4 hrs after induction with 0.5 mM IPTG. The cell lysis was performed using the same protocol of the pro-cat MMP-1.

The protein was purified by the ion exchange column (Hitrap Q, Pharmacia), using a linear gradient of NaCl up to 0.35 M in 6 M urea, 20 mM Tris (pH8.0). The refolding of the protein was followed as described in below.

The purified and denatured protein was diluted in a 4 M urea solution containing 50 mM Tris (pH 7.4), 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl, keeping the protein concentration below 0.15 mg/ml to minimize intermolecular interactions which could cause misfolding. During the dialysis steps, the concentration of urea was gradually decreased from 4 M to 2 M and then completely removed from the buffer. The presence of self-hydrolysis has to be taking into account during the refolding process. For this reason, starting from the 2 M urea step, 0.2 M of acetohydroxamic acid (AHA), a mild MMPs inhibitor, was added to the refolding buffers. The choice of a mild inhibitor is of

foundamental importance since it can be easily removed or replaced by a stronger one. The folding of the protein has been investigated by NMR through ^{1}H-^{15}N HSQC (Heteronuclear single quantum coherence) spectra. A good spreading of the cross-peaks was found. The activity has been checked using a fluorescent substrate or a reactive thiolic peptide (see in Methods). The biophysical and biological properties of the refolded protein are summarized in Table 2 at the end of this chapter.

Thanks to the developed methods for expression, purification, and refolding of the catalytic domain of MMP-1, large amounts of protein were produced and employed to generate protein-ligand structural models by a combination of *in silico* tools and experimental NMR data (Chapter 4.1). The detailed results are described in Publications.

MMP-7

The cDNAs encoding the sequence for the catalytic domain of MMP-7 (Tyr95-Lys267) was cloned into pET21 expression vector using *Nde I* and *Hind III* (for MMP-7) as restriction enzymes and expressed in BL21(DE3) cells. Test expression of the protein was performed in minimal media at 37 ^{0}C and SDS-PAGE analysis of the test expression is shown in Figure 8.

The cell lysis was performed using the same protocol of the pro-cat MMP-1 with an additional step consisting in a dialysis against a 2 M urea solution containing 2% Triton, 50mM Tris (pH 8.0), 5mM EDTA, 1mM DTT in order to remove the impurities. Then the protein was washed three times by 50mM Tris (pH 8.0), 5mM EDTA, 1mM DTT. After cell lysis the inclusion bodies were solubilized in 8 M urea solution containing 20 mM Tris pH (8.6). The protein was purified by the ion exchange column (Hitrap Q) using a linear gradient of NaCl up to 0.35 M with a buffer containing 20 mM Tris pH (8.6) and 6 M urea.

However the previous refolding methods used for MMP-1 were not suitable for this protein. Also by using the pro-cat MMP-7 construct (Leu20-Lys267), disappointing

results were obtained in terms of folding and solubility. These problems have been partially overcome using a new refolding strategy developed during this PhD project.

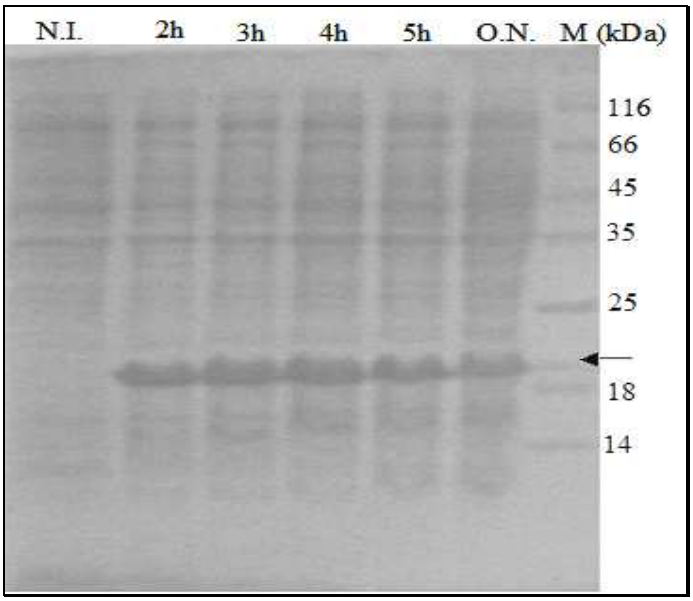


Figure 8. SDS-PAGE analysis of test expression of the catalytic domain of MMP-7 in BL21(DE3) cells at 37 ^{0}C

The purified MMP-7 was diluted (<0.15mg/ml) in 4 M urea, 50 mM MES (pH 6.5), 1mM $ZnCl_2$, 10 mM $CaCl_2$, 0.3 M NaCl at 4 ^{0}C. The diluted protein was dialyzed against 4 M urea, 50 mM (Tris pH 7.0), 1mM $ZnCl_2$, 10 mM $CaCl_2$, 0.3 M NaCl. Then the concentration of urea was gradually decreased from 4M to 2M and finally completely removed from the buffer. The refolded protein was dialyzed against 10 mM Tris (pH 7.2), 0.1 mM ZnCl2, 10 mM $CaCl_2$, 0.3 M NaCl, 0.2 M AHA and then concentrated at 4 ^{0}C using an Amicon stirrer, fitted with a YM10 membrane in nitrogen atmosphere and with Centriprep concentrator, fitted with YM10 membranes. The final protein samples showed a good spreading of the ^{1}H-^{15}N HSQC spectra. The biophysical and biological properties of the protein are summarized in Table 2 at the end of this chapter.

MMP-8

The catalytic domain of MMP-8 (Asn85-Gly242, numbering from ref. 65) was cloned into the pET21 expression vector using *Nde I* and *Xho I* as restriction enzymes and expressed in BL21(DE3) Codon Plus cells. For the ^{15}N- and ^{13}C-labeled protein, the test expression was performed in minimal media at 37 ^{0}C (Figure 9). After induction with 0.5 mM IPTG the best expression was obtained in 5hrs.

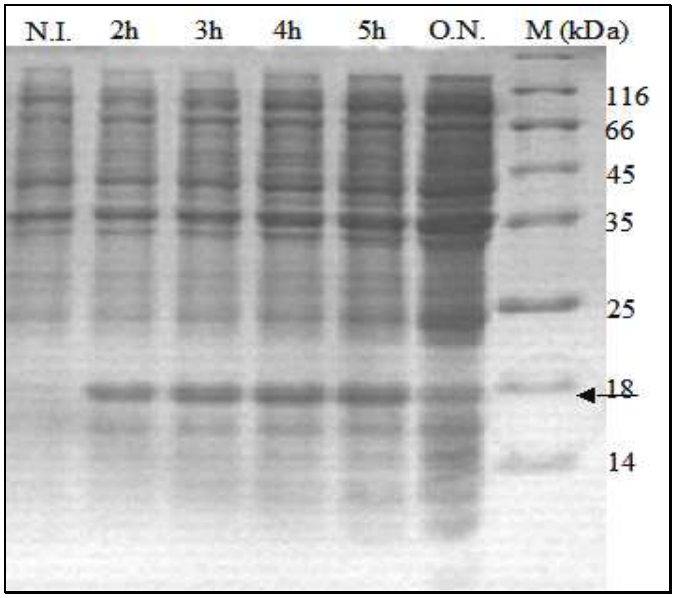


Figure 9. SDS-PAGE analysis of test expression of the catalytic domain of MMP-8 in BL21(DE3) codon plus cells at 37 ^{0}C

The cell lysis was performed using the same protocol for the MMP-7. After cell lysis the inclusion bodies were solubilized in 8 M urea solution containing 20 mM Tris (pH 8.0) and then purified by the Hitrap Q column using a linear gradient of NaCl up to 0.35M. The refolding of the protein was performed as described for the catalytic domain of MMP-1 (Asn106-Gly261). By using this protocol the protein showed a good spreading of ^{1}H-^{15}N HSQC spectra. The biophysical and biological properties of the protein are summarized in Table 2 at the end of this chapter. These protocols allowed us to produce large amounts of stable and soluble samples which have been used in structural studies by NMR and x-ray

crystallography.

MMP-12

MMP-12 (Gly106-Gly263) was cloned into pET21 expression vector and expressed in BL21(DE3) [66]. For inactive form (Glu219Ala), a single point mutation was performed using the QuickChange™ Site-Directed Mutagenesis Kit. The test expression of the protein was performed in minimal media at 37 ^{0}C (Figure 10).

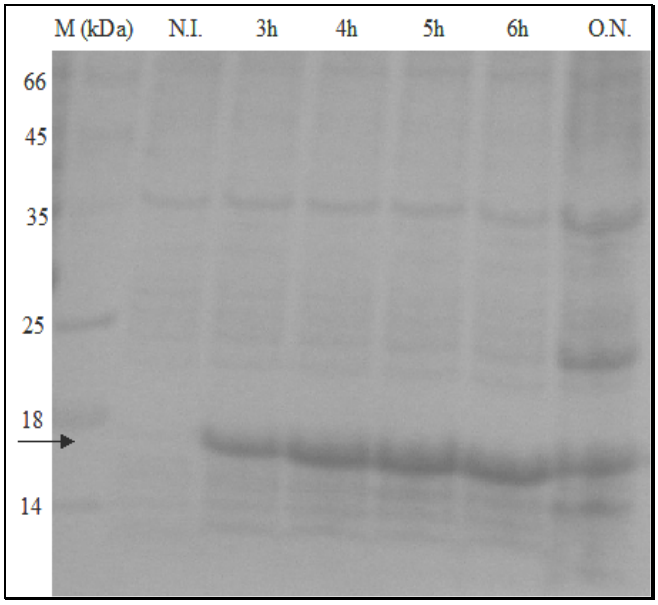


Figure 10. SDS-PAGE analysis of test expression of the catalytic domain of MMP-8 in BL21(DE3) cells at 37 ^{0}C.

The cell lysis was performed using the protocol already developed for the MMP-7. After cell lysis, the inclusion bodies were solubilized in a buffer containing 6 M urea and 20 mM sodium acetate (pH 5.0). The protein was purified by the ion exchange column, Hitrap SP (Pharmacia), using a linear gradient of NaCl up to 0.35M.

For the refolding of MMP-12 (Glu219Ala), the methods previously described were

not suitable. Despite many attempts using different conditions such as a slow decreasing of urea concentration, direct refolding to 0 M urea, different concentration of NaCl or $CaCl_2$, unfolded protein was obtained. Finally, we found that high concentration $ZnCl_2$ (1 mM) like for MMP-7 was needed for the refolding. The concentration of $ZnCl_2$ is then decreased up to 0.1 mM in the last dialysis. Therefore, for proteins such as MMP-7 and MMP-12 (Glu219A) an high concentration of $ZnCl_2$ is decisive for a good refolding. Although this phenomenon is unknown one hypothesis is that the high concentration of $ZnCl_2$ may favor a correct fold of the denatured proteins by interacting with the histidine ligands. The biophysical and biological properties of the refolded proteins are summarized in Table 2. The good quality of the protocols developed is demonstrated by the several biophysical investigations carried out on MMP-12 in this laboratory during this PhD project.

The folding of all the proteins has been investigated by NMR through ^{1}H-^{15}N HSQC spectra. The activity, checked using a fluorescent substrate or a reactive thiolic peptide, was in agreement with the already published data. For all the proteins a good spreading of the resonances was found. With the exception of MMP-8, all proteins exhibited a good stability in presence of the weak inhibitor AHA. In the case of the MMP-8 catalytic domain, even in the presence of AHA, a modest self-hydrolysis was still present at concentrations above 0.5 mM.

3.2 Crystallization procedures

Even if the protein exhibits limited stability, crystals of the catalytic domain of MMP-8 have been obtained by growing them at 20° C in a solution containing 0.1 M Tris-HCl (pH 8.0), 20% PEG-3350, 0.2 M $MgCl_2$. The final concentration was around 8 mg/ml of protein. The crystals of the catalytic domain of MMP-12 were obtained as previously reported [66].

Table 2. Biophysical and biological properties of the catalytic domain of MMP-1, MMP-7, MMP-8, MMP-12. *The result from our group [66]. [a]The activity of the catalytic domain of human MMP-12 (wild type) in the same assay condition at 37 ^{0}C (BIOMOL, Catalog No: SE-138)

Properties	MMP-1 (Asn106-Gly261)		MMP-7 (Tyr95-Lys267)	MMP-8 (Asn85-Gly242)	MMP-12 (Gly106-Gly263)	
	Wild type	E219A mutant (inactive)			F171D mutant* (active)	F171D, E219A mutant (inactive)
Yield (mg/L) from minimal media	~8-10	~8-10	~15-18	~12-15	~30-40	~30-40
Purification	Ion exchange	Ion exchange	Ion exchange	Ion exchange	Ion exchange	Ion exchange
Yield (%) after refolding	>80	>50	<30	>75	>80	>70
Folding state	Good	Good	Good	Good	Good	Good
Solubility (mM)	>2.0 mM	>2.0 mM	>0.5 mM	>1.2 mM	>2.0 mM	>2.0 mM
Stability of 0.5 mM protein at R.T. for a week containing 0.5 M AHA	Stable	Stable	Stable	Stable in <0.5 mM	Stable	Stable
Activity (U/μg)	686	-	23	223	91 (86[a])	-

In both cases the crystallization buffer contained 0.2 M of the weak inhibitor (AHA). Crystals of MMP12-AHA were obtained in the presence of $LiCl_2$ as well. The complexes were obtained through soaking of MMP12-AHA crystals with a solution containing the inhibitor. To obtain the active uninhibited enzymes, MMP crystals were extensively dialyzed against the same crystallization buffers lacking AHA.

In order to deeply characterize the different steps of the catalytic mechanism of the

MMPs, adducts of MMP-12 and MMP-8 with the collagen fragment ProGlnGlyIleAlaGly, which is known to be cleaved at the Gly-Ile bond, were prepared. As expected, the catalytic activity of the crystallized proteins results in a selective hydrolysis of the peptide. The peptide was soaked into the crystals for 1-3 days in order to obtain the two- or the one-peptide adducts. Considering that the protein tends to digest itself, the crystallization was always carried out in the presence of AHA inhibitor. By playing with various crystals under crystal washing conditions we got the crystals of the uninhibited form of MMP-12 and of MMP-8. Also performing the experiments with the Glu219Ala mutant of MMP-12, no peptide-bound form could be identified, due to rapid hydrolysis. Therefore also the Glu219Ala mutant of MMP-12, whose MMP-1 analogue is known to retain only 0.1% of the enzyme activity, apparently hydrolyses the substrate [67]. On the other hand, well-resolved structures with a hydrolysis product inside the active site have been obtained for both MMP-12. All the adducts mentioned above have been solved by x-ray crystallography. The detailed results are described in the attached publications.

3.3 Expression of the catalytic and hemopexin domain (full length) of MMP-12 and the hemopexin domain of MMP-12

Our data shows that the full length active MMP-12 is highly sensitive to the self-hydrolysis. In fact the protein is already completely degraded during the refolding steps. To avoid this phenomenon, the mutant Glu219Ala has been cloned and expressed. On the contrary the hemopexin domain of MMP-12 is stable during purification and refolding since the catalytic domain is absent.

The test expression for the isotope labeled proteins was performed in minimal media at 37 ^{0}C (Figure 11). The active full length MMP-12 (Glu219Ala) was expressed in BL21 Codon Plus strains overnight after induction with 0.5 mM IPTG, and the hemopexin domain of MMP-12 was expressed in BL21(DE3) strains for 5 hrs after induction with 0.5 mM IPTG.

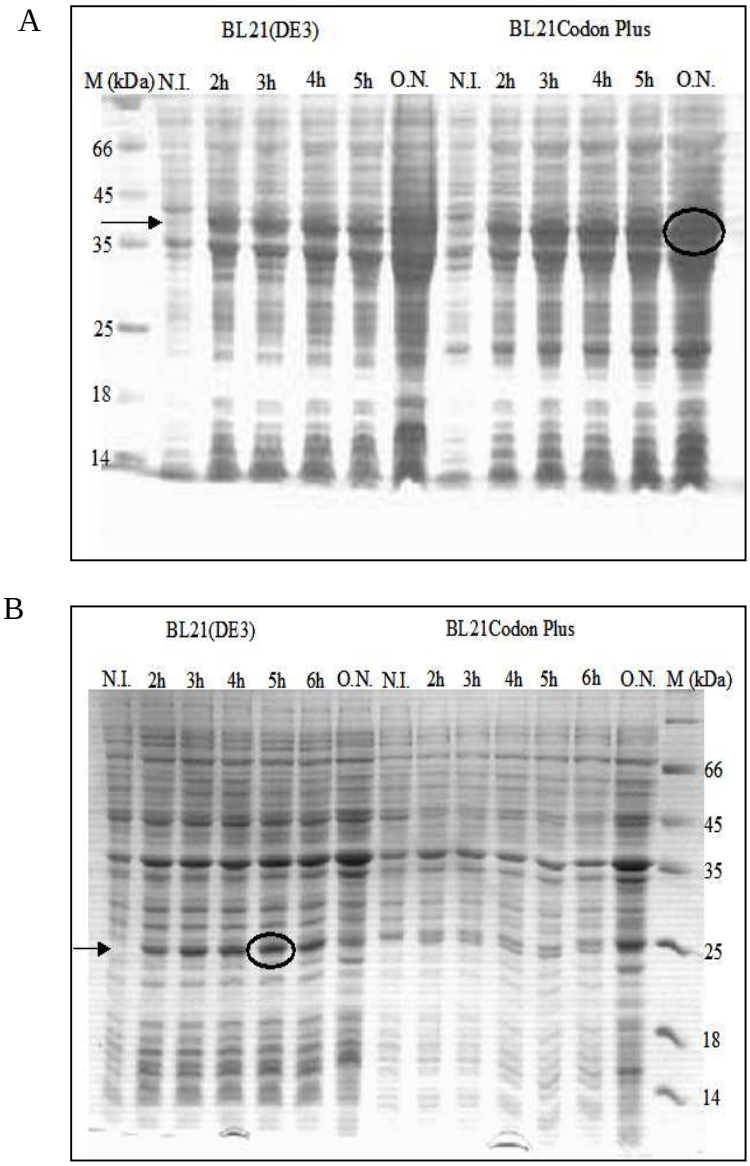


Figure 11. SDS-PAGE analysis of test expression of the full length (A) and the hemopexin domain (B) of MMP-12 in minimal media at 37 ^{0}C

3.4 Development of purification and refolding methods for the full length of MMP-12 and the hemopexin domain of MMP-12

Considering that both the full length protein and the hemopexin domain construct contain two cysteines, a new refolding protocol has been developed. To avoid oligomerization phenomena, a solution of 6 M guanidinium chloride with reductants was

used to dissolve the inclusion bodies. Then the denatured protein was directly refolded without any purification by multi-step dialysis. However this approach was not efficient since an extensive precipitation occurred.

In the new approach the detergent Brj-35 is used to avoid the precipitation of the protein during the refolding. However, the detergent represents an obstacle for structural studies and it is not easy to remove. During the project we found that the detergent could be easily removed from the buffer, by using a weak cation exchange column (HiTrap CM FF, Pharmacia) and a gel filtration column (HiLoad™ 16/60: 120 ml, Pharmacia). The detailed protocol is described below.

The cell lysis was performed using the same protocol of the pro-cat MMP-1 with an additional washing by a 2M urea solution containing 2% Triton, 50mM Tris (pH 8.0), 5mM EDTA, 1mM DTT. Then the sample was further washed twice with 6 M urea, 50 mM Tris (pH 8.0). Most of the impurities were removed during cell lysis by washing the sample with a 6 M urea solution. The inclusion bodies were completely denatured using a solution containing 6 M guanidinium chloride, 50 mM Tris (pH 8.0), 10 mM dithiothreitol (DTT), 20 mM cystamine at 4 ^{0}C overnight. Then the supernatant was collected after centrifugation. The concentration of the denatured protein was roughly measured using the calculated extinction coefficient of the protein at 280 nm. The denatured protein was then diluted (< 0.15 mg/ml) in 6 M guanidinium chloride, 20 mM Tris (pH 8.0), 1 mM DTT, 0.05% Brij-35 and dialyzed against 20 mM Tris pH 7.2, 0.1 mM $ZnCl_2$, 10 mM $CaCl_2$, 150 mM NaCl, 0.05% Brij-35, 5 mM β-mercaptoethanol, 1 mM 2-hydroyethyl disulfide. In the last dialysis the sample was dialyzed against the same buffer without β-mercaptoethanol. To remove the excess of oxidant and Brij-35, the protein was dialyzed against a solution containing 20 mM Tris (pH 7.2), 0.1 mM $ZnCl_2$, 10 mM $CaCl_2$, 150 mM NaCl and then several times against a buffer containing 20 mM Tris (pH 7.2), 10 mM $CaCl_2$. Then the refolded protein was centrifuged to remove the precipitate, and purified on 5 ml (x 2) of HiTrap CM FF column

(Pharmacia) using 400 ml (20 CV) of 20 mM Tris (pH 7.2), 10 mM $CaCl_2$. The protein was then eluted using the same buffer containing 0.3 M NaCl and 0.1 M AHA. To further purify the protein a gel filtration was performed using the same buffer. For NMR studies, the purified protein was washed with a buffer solution containing 20 mM Tris (pH 7.2), 10 mM $CaCl_2$, 0.3 M NaCl, 0.2 M AHA, and 3 mM of the nanomolar inhibitor NNGH. The same protocol with minimal modification has been used for to refold the hemopexin domain of MMP-12. The formation of disulfide bridge between the two cysteins of the hemopexin domain has been checked by SDS-PAGE (Figure 12). Under reducing conditions, a 42 kDa band (for the construct Gly106-Cys470) corresponding to the reduced form of the full length of MMP-12 (Figure 12. A) and a 23 kDa band (for the construct Glu278-Cys470) corresponding to the reduced hemopexin domain of MMP-12 (Figure 12. B) were observed. On the contrary under non-reducing conditions a single dominant band appeared below the reduced form of each protein.

This refolding protocol has provided protein samples with a well spreading of the cross peaks in the ^{1}H-^{15}N TROSY (transverse relaxation-optimized spectroscopy) spectra (Figure 13). The solution structure the hemopexin domain of MMP-12 is in progress.

However the “inactive” mutant of MMP-12 full length was still sensitive to self-digestion even in presence of inhibitors and it is rapidly degraded at concentration above 0.3 mM. The cleavage site has been identified performing the mass spectra of the resultant fragments. In particular the cleavage occurs in the linker (hinge region) between the amino acids Glu267 and Asn268. The sensitivity of the active full length MMPs to the self-hydrolysis has been already reported [41, 43, 63, 68-71]. At this regard, several researchers believe that this process may have an important physiological role [68-69]. The tendency to the self-hydrolysis can be strongly reduced by replacing the catalytic zinc ion with Cd^{2+} without relevant structural alteration [72]. The metal replacement was obtained by performing a multi-step dialysis against buffer solutions containing Cd^{2+} instead of Zn^{2+}. The activity of

the full length Cd^{2+}-MMP-12 was 400-fold lower with respect to that of the full length Zn^{2+}-MMP-12 protein. Although a very slow self-hydrolysis is still present, nevertheless the Cd^{2+}-MMP-12 protein is much more stable with respect to the zinc analog allowing us to perform NMR experiments for the determination of the structure in solution. The solution structure of the protein is in progress. To further stabilize the protein we have already planned the expression of new mutants at the hinge region.

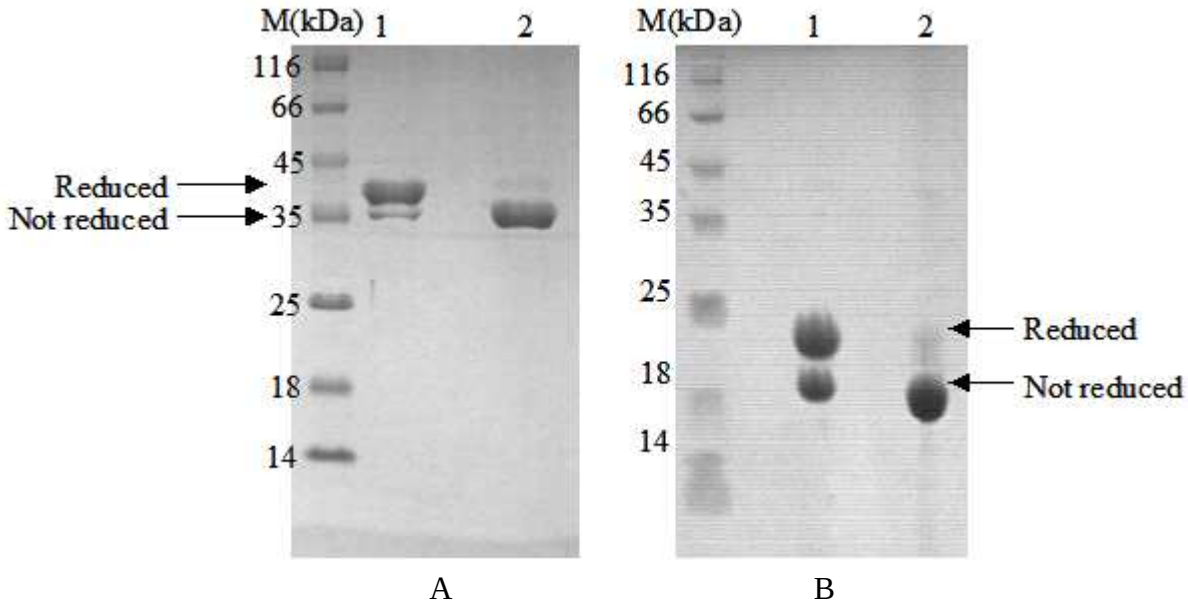
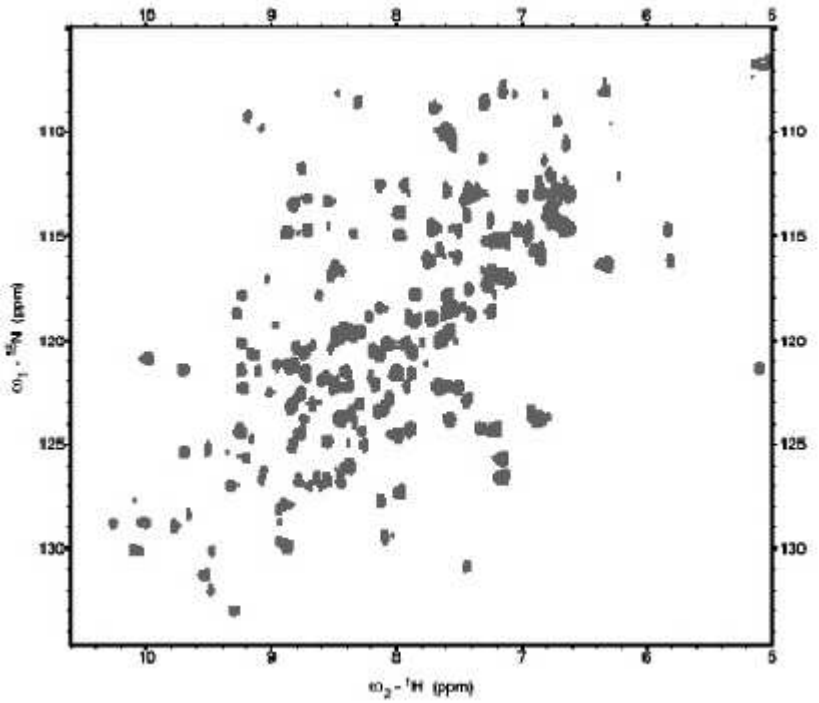


Figure 12. SDS-PAGE analysis of the disulfide bridge of the full length of MMP-12 (A) and the hemopexin domain of MMP-12 (B) at reducing and non-reducing conditions. M, protein marker; lane 1, with 50 mM DTT; lane 2, without DTT.

A. The hemopexin domain of MMP-12



B. The full length of MMP-12

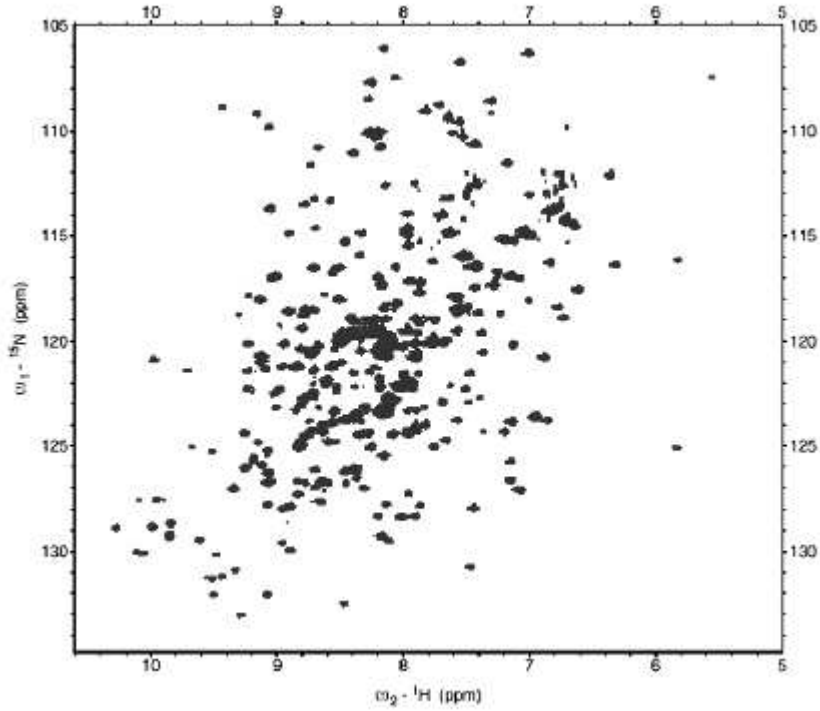


Figure 13. ^{1}H-^{15}N TROSY spectra of the hemopexin domain (A) and the full length (B) of MMP-12

4 Publications

4.1 Combining *in Silico* Tools and NMR Data To Validate Protein-Ligand Structural Models: Application to Matrix Metalloproteinases [*J. Med. Chem*., 2005, 48, 7544-7559]

Combining in Silico Tools and NMR Data To Validate Protein−Ligand Structural Models: Application to Matrix Metalloproteinases

Ivano Bertini,*[,†,‡] Marco Fragai,[†,§] Andrea Giachetti,[†,⊥] Claudio Luchinat,[†,§] Massimiliano Maletta,[†,‡] Giacomo Parigi,[†,§] and Kwon Joo Yeo[†]

Magnetic Resonance Center, University of Florence, Via Luigi 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 Donizetti 6, 50144 Florence, Italy, and ProtEra, Viale delle Idee 22, 50019 Sesto Fiorentino, Italy

Received June 17, 2005

A combination of in silico tools and experimental NMR data is proposed for relatively fast determination of protein−ligand structural models and demonstrated from known inhibitors of matrix metalloproteinases (MMP). The ^{15}N ^{1}H heteronuclear single quantum coherence (HSQC) spectral assignment and the 3D structure, either X-ray or NMR, are needed. In this method, the HSQC spectrum with or without the ligand is used to determine the interaction region of the ligand. Docking calculations are then performed to obtain a set of structural models. From the latter, the nuclear Overhauser effects (NOEs) between the ligand and the protein can be predicted. Guided by these predictions, a number of NOEs can be detected and assigned through a HSQC NOESY experiment. These data are used as structural restraints to reject/refine the initial structural models through further in silico work. For a test protein (MMP-12, human macrophage metalloelastase), a final structure of a protein−ligand adduct was obtained that matches well with the full structural determination. A number of structural predictions were then made for adducts of a similar protein (MMP-1, human fibroblast collagenase) with the same and different ligands. The quality of the final results depended on the type and number of experimental NOEs, but in all cases, a well-defined ligand conformation in the protein binding site was obtained. This protocol is proposed as a viable alternative to the many approaches described in the literature.

Introduction

Rational drug design strategies must rely on the availability of high-throughput methods to experimentally determine the structure of candidate drug-target complexes.[1] The obtained structural information is then used to improve and optimize the candidate drug in a cyclic procedure. Obtaining three-dimensional macromolecular structures is still a time-consuming task. X-ray structure determination is becoming a high-throughput method,[2] but the method requires the easy availability of protein crystals that are suitable for soaking with the various candidate drugs. NMR is also a high-throughput technique in drug discovery,[3,4] but its power lies mostly in the earlier phases of the process, i.e., in the first screening of a relatively large number of compounds. NMR quickly provides information on binding affinity and on the region of interaction of the candidate drug with the target molecule.[5]

NMR is of course also able to determine the three-dimensional structure of the adduct, but the procedure is time-consuming.[6] Moreover, obtaining a 3D structure depends on the full assignment of thousands of intraprotein nuclear Overhauser effect spectroscopy (NOESY) cross-peaks, while the only relevant ones are the few intermolecular cross-peaks between protein and ligand signals. In silico prediction of the structure of the adduct through docking programs, while valuable in the early ligand design phases, is not reliable at this stage.[7−9] Independently of the docking program used, in many cases more than one binding poses are found that do not significantly differ in predicted binding energies.

The availability of a fast and reliable method able to provide a molecular model based on few experimental restraints is an ambitious goal for overcoming these problems. Recently, several efforts have been performed in this direction.[10−13] For instance, a suite of NMR experiments has been recently proposed as a tool to provide structural information on protein−ligand adducts,[12] through intermolecular NOEs detected in selectively labeled proteins. The method is applicable to very large proteins once their three-dimensional structure is known.

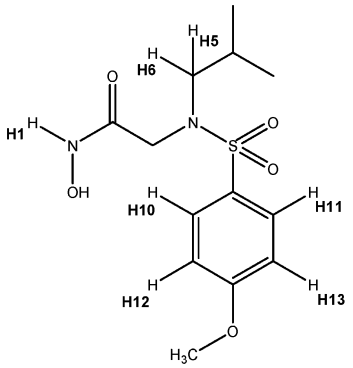
For smaller proteins, it is worth to investigate whether a few NOEs may be obtained even without selective labeling of the proteins. We propose here a combined use of computational tools and a small number of experimental NMR restraints as an efficient way of selecting the correct binding pose among those proposed by docking programs. The experimental restraints are (i) the heteronuclear single quantum coherence (HSQC) chemical shifts to select the region of interest on the

* To whom correspondence should be addressed. Tel: +39-055-4574270. Fax: +39-055-4574271. E-mail: bertini@cerm.unifi.it.
† Magnetic Resonance Center, University of Florence.
‡ Department of Chemistry, University of Florence.
§ Department of Agricultural Biotechnology, University of Florence.
⊥ ProtEra.

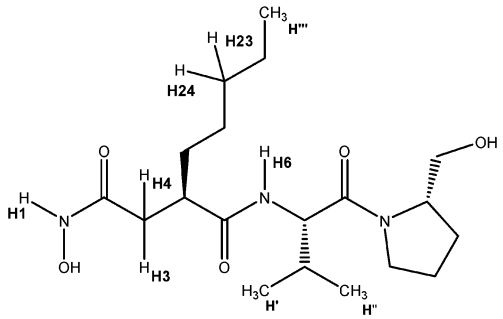
10.1021/jm050574k CCC: $30.25
Published on Web 11/05/2005

Chart 1

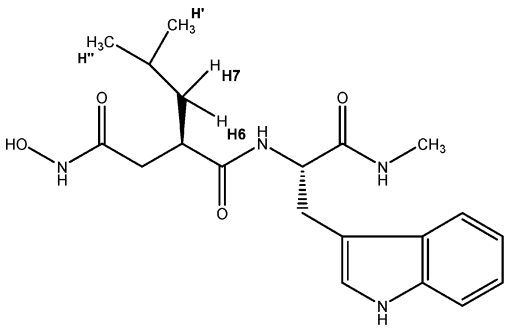
NNGH



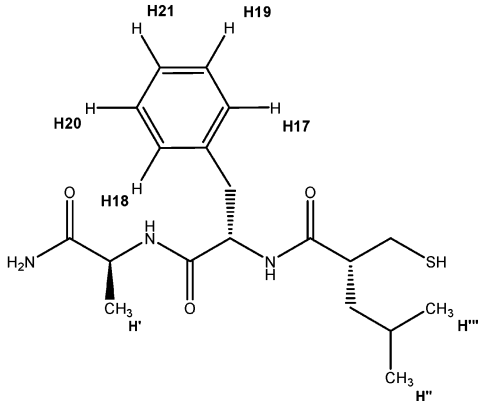
Actinonin



Galardin



SIMP-1



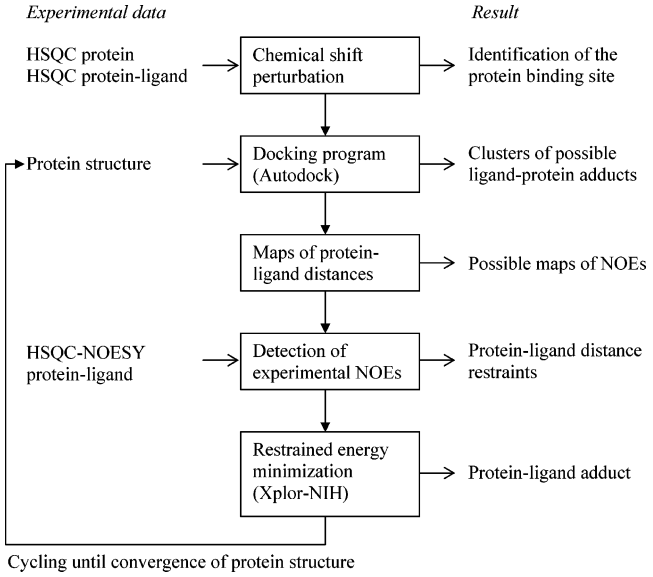
target, and (ii) the few ligand-target NOEs that can be unambiguously identified from ^{15}N NOESY−HSQC experiments. Besides the protein three-dimensional structure, only a singly ^{15}N-labeled protein sample and a preexisting assignment of its ^{15}N ^{1}H HSQC spectrum are required.

The method has been validated by reproducing the known docked conformation of *N*-isobutyl-*N*-[4-methoxyphenylsulfonyl]glycyl hydroxamic acid (NNGH, see Chart 1) bound to matrix metalloproteinase 12 (MMP-12, human fibroblast metalloelastase). The method has been then applied to obtain the docked conformations of NNGH and other three ligands, (3-[[1-[[2-(hydroxymethyl)-1-pyrrolidinyl]carbonyl]-2-methylpropyl]carbamoyl]-octanohydroxamic acid (actinonin), *N*-[(2*R*)-2-(hydroxamidocarbonylmethyl)-4-methylpentanoyl]-L-tryptophan methylamide (galardin), and (2*R*)-2-mercaptomethyl-4-methylpentanoyl-L-phenylalanyl-L-alanine amide (SIMP-1) (see Chart 1) to MMP-1 (human fibroblast collagenase). MMPs belong to a family of zinc-dependent endopeptidases responsible for the metabolism of extracellular matrix proteins,[14−16] and alterations in their levels are implicated in a wide range of pathological states,[17,18] so that these proteins represent attractive drug targets.

Methods

The protocol consists of the following steps, reported in Scheme 1: (a) identification of the protein binding site, (b) calculation of possible protein−ligand adducts, (c) prediction of the map of NOEs corresponding to each computed conformation, (d) determination of few experimental restraints, able to select the real adduct among those calculated, and (e) validation and cyclic in silico refinement of the ligand position in the protein scaffold. The identification of the protein binding site can be conveniently performed from the analysis of the chemical shifts acquired in the presence and in the absence of the ligand. NOEs between ligand protons and protein protons are obtained from ^{15}N NOESY−HSQC

Scheme 1



spectra. The protocol requires that the protein structure and the assignment of its ^{15}N ^{1}H HSQC spectrum is known.

HSQC spectra of the protein in the presence and in the absence of the ligand must be acquired. Most of the protein peaks will coincide in the two spectra. Only peaks corresponding to amide protein protons close to the ligand will be in different positions, but their shift is usually small enough to be easily assigned. This information is used to identify the protein binding site, according to the value of the combined $^{1}H/^{15}N$ shift perturbation upon complexation, given by $\Delta = (\Delta\delta(^{1}H)^2 + (\Delta\delta(^{15}N)/6)^2)^{1/2}$.[19] The residues with a significantly large value of Δ, except those at sizably larger distance from all others, are used to identify the grid for docking calculations. The latter is centered on the protein surface atom closest to the center of the smallest sphere that comprises all the selected nitrogen atoms.

Due to the complexity of the energy landscape on the path to the global minimum region,[20] a specific ligand−protein docking program is invoked in order to accurately probe and select the conformations of the ligand according to appropriate scoring functions. We use the program Autodock because it has been amply validated and tested on the target proteins selected for this study. The docking program can be run to obtain clusters of the possible adducts. Such clusters are then used to predict NOEs between protein and ligand nuclei. In fact, a map of distances between ligand and protein nuclei can be obtained for each of the different clusters. The presence of cross-peaks can thus be predicted for the different possible adducts and compared with the cross-peaks actually present in the experimental spectra.

The following experiments must be performed: ^{15}N NOESY−HSQC spectra of the protein−ligand adduct and of the free protein, and the 1D ^{1}H spectrum of the free ligand in water. The latter experiment provides an estimate of where the chemical shifts of ligand signals in the adduct have to be looked for. The presence of intermolecular cross-peaks, i.e., peaks between frequencies close to those of the free ligand in one dimension, and those of the protein amide protons predicted to be in the vicinity of the ligand in the other dimension, is checked. Such cross-peaks, if absent in the free protein spectrum and not attributable to nuclei of other neighboring protein residues, are unambiguously assigned. A good correspondence between expected and observed cross-peaks is a clear indication of the goodness of the corresponding cluster. On the other hand, direct evidence of the unacceptability of some of the clusters generated by Autodock can be obtained. The experimental NOEs, translated into upper distance limits, can then be used to refine the remaining acceptable structures and possibly to further discriminate among them. The refinement procedure has been developed using Xplor-NIH. In such procedure, the protein side chains are left free to move, thus allowing a better docking to be obtained with respect to docking programs where the protein is completely rigid.

The refinement procedure consists of loading the calculated adduct and performing an in vacuo molecular dynamics simulation in internal coordinates, with backbone atoms grouped together to constitute a rigid structure. A simulated annealing is performed by heating the system to 1500 K and then cooling it to 50 K in steps of 50 K. At each temperature, 750 steps of molecular dynamics simulations are performed with time steps of 2 fs. The force constant of NOE restraints is fixed to 30 kcal mol^{-1} $Å^{-2}$, and van der Waals, electrostatic terms and the protein and ligand force field (angles, bonds, dihedrals and impropers) are also included. The resulting structures are then refined with a Powell minimization, and ordered according to the value of the target function. The latter is calculated considering the ligand-residue and residue−residue interactions only for residues up to 8 Å from the ligand. This helps reducing the energy "noise" originating from slight changes in residue−residue interactions far away from the ligand site. In all cases, the best 10 structures over 200 calculated through Xplor-NIH starting from each tentative docking structure are very similar to one another.

The structure of the adduct is thus calculated through the consecutive use of the program Autodock and the refinement procedure working in Xplor-NIH. Xplor-NIH calculations can significantly change the protein side chain positions after complexation. Therefore, cycling between Autodock and Xplor-NIH refinement is necessary until convergence to a fixed protein structure is achieved. We have tested that such approach can actually select the correct ligand−protein docking, among those proposed by Autodock. Furthermore, the introduction of experimental data and the allowed mobility of the protein side chains provide more confidence in the obtained adduct.

MMP systems, the receptors that we used in this work, have a catalytic zinc ion as active center, coordinated to three histidines. The three zinc-coordinated histidines were treated as the neutral form with the hydrogen on ND1, whereas other histidines used the default option with hydrogen on NE2. Glutamates were treated as charged form as default, except the catalytically essential glutamate 219,[21] at the second shell of the zinc binding site. The latter residue was protonated, with the hydrogen on the oxygen nearest to the catalytic zinc, or deprotonated depending on whether the zinc donor atom closest to it was deprotonated (hydroxamate ligands)[21] or protonated (thiol ligands). To take into account the electron density delocalization due to coordination of ligands, the charge of the zinc ion was distributed among the protein ligands.[21]

Results

Test with a Known Structure: MMP-12−NNGH. NNGH is a broad spectrum MMP inhibitor able to interact with both the catalytic zinc and the S1′ cavity.[6,22] In particular, it is able to bind MMP-12 with nanomolar affinity (K_d = 10 nM),[6] and for this reason it has been chosen as a model system to study protein-inhibitor interactions. Its molecular structure is reported in Chart 1.

The structure of MMP-12 complexed to NNGH is already known,[6] as both the X-ray (Figure 1A) and the NMR structures of the adduct have been solved. Therefore, we used such system as a test for our protocol. The HSQC spectra of the protein without and with the NNGH in solution were acquired. Inspection of residues showing significant chemical shift perturbation (see

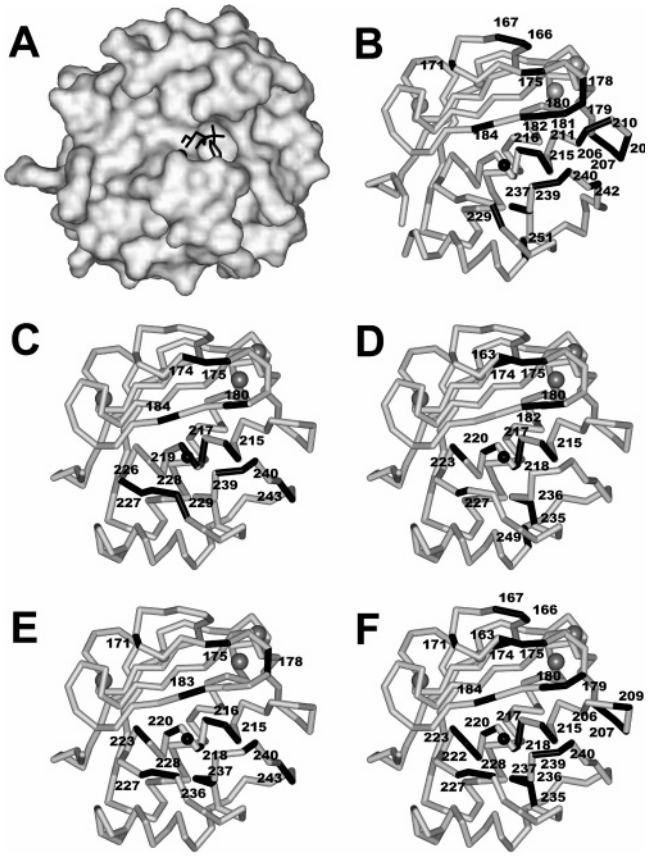


Figure 1. X-ray structure of the MMP-12−NNGH adduct (PDB 1RMZ) (A), residues of MMP-12 affected by chemical shift perturbation upon complexation with NNGH (B), and residues of MMP-1 affected by chemical shift perturbation upon complexation with NNGH (C), actinonin (D), galardin (E), and SIMP-1 (F) (see Table 2).

Table 1) permitted to define the ligand binding region as the protein catalytic site (defined here as constituted by the zinc binding region, the S1′ pocket and the substrate binding groove) with reasonable accuracy. As expected from the crystal structure of the MMP-12−NNGH adduct, among the affected resonances are residues 210, 211, 215, and 216 on the α helix at the bottom of zinc binding site, residues 237, 239−240 and 242 forming the hydrophobic S1′ cavity, and residues 179−182 and 184 on the strand facing both the catalytic metal and the S1′ pocket (Figure 1B).

Calculations were performed using the X-ray structure of the protein (PDB 1Y93) at 1.03 Å resolution.[6] Autodock was used to select the lower docking energy conformations. Docked conformations were clustered according to a maximal RMSD of 1 Å (Figure 2). The docking energies for the first, second, third and fourth clusters were −15.89, −15.84, −15.01, and −14.44 kcal mol^{-1}, respectively. The second cluster is in accordance with the X-ray structure of the adduct (PDB 1RMZ).[6] The plane containing the hydroxamic group in the first and third cluster is oriented perpendicularly to the plane containing the hydroxamic group in the second cluster. The *p*-methoxy-phenyl group enters more deeply in the S1′ pocket in the first than in the third cluster. In the fourth cluster the *p*-methoxy-phenyl group does not sit in the S1′ pocket.

These structures were separately refined with Xplor-NIH using the already available NOEs with protein backbone NH atoms (see Figure 3B).[6] The second cluster remains essentially unchanged, with total energy −1184 kcal mol^{-1} (see Figure 2). The structure calculated using the third cluster as starting conformation is similar to the previous one, with total energy −1179 kcal mol^{-1}. The structure calculated from the first cluster has total energy −1062 kcal mol^{-1}, and no coordination of the hydroxamic group to the metal ion; the one calculated from the fourth cluster has total energy −734 kcal mol^{-1}, and the *p*-methoxy-phenyl group outside the S1′ pocket.

Slight changes in the side-chain protein structure were obtained, and new Autodock calculations were thus performed using the three lowest energy Xplor-NIH protein structures. Remarkably, the lowest docking energy clusters calculated by Autodock now converge to similar conformations using the second and third Xplor-NIH protein structure (see Figure 2). These conformations are in agreement with the X-ray structure, with docking energy from −15.85 kcal mol^{-1} to −15.63 kcal mol^{-1}. Xplor-NIH refinements provided structures (see Figures 2 and 3B) with lowest total energy from −1197 to −1180 kcal mol^{-1}, in agreement with the X-ray structure (see Figure 3A).

Analogous calculations were performed also using the X-ray structure PDB 1OS9, with 1.85 Å resolution.[23] In this structure the active site of one molecule is not hosting an external ligand but the N-terminal part of the neighboring protein molecule. The calculations converged to the same adduct obtained starting from the 1Y93 structure.

Determination of Structural Models for Ligand Adducts of MMP-1. NNGH itself and three other known strong inhibitors of MMPs were selected as representatives of different classes of ligands and tested against MMP-1. The test consists in following the protocol described above and checking whether (i) unambiguous NOEs could be obtained and (ii) the cycling between Autodock and Xplor-NIH calculations permits the selection of one ligand conformation. Calculations were performed using the X-ray structure of the inhibitor-free protein (PDB 1CGE) with 1.90 Å resolution.[24]

MMP-1−NNGH. The first ligand examined is the same ligand used to validate the protocol with MMP-12. The structure of the NNGH adduct with MMP-1 is not known, although it is reasonable to believe that it will adopt a similar conformation. We measured an IC_{50} value for the adduct of 174 nM.

Chemical shift perturbation affects the zinc binding histidine 228 and the neighboring residues 226, 227, and 229, residues 239, 240, and 243 forming the S1′ hydrophobic pocket, residues 215, 217, and 219 on the α-helix where the metal binding site is inserted, and residues 180 and 184 on the parallel strand (see Figure 1C). This is an expected feature, but it is a new independent experimental information based on which an Autodock grid was generated. The grid resulted nicely centered around the known catalytic site. Autodock calculations using this grid were thus performed.

The four lower docking energy clusters were analyzed. The docking energies were −14.68, −13.99, −13.94, and −13.81 kcal mol^{-1}, respectively. The structures in the first and third clusters show similar hydroxamate

Table 1. MMP Residues Subjected to Significant Chemical Shift Perturbations (in bold)[a]

seq	163									171	..	174											186	..	204					209
MMP12+NNGH	F	A	R	**G**	**A**	*H*	*G*	D	**D**	..	F	**D**	G	*K*	**G**	**G**	**I**	**L**	**A**	H	**A**	F	G	..	T	T	**H**	**S**	**G**	G
MMP1+NNGH	F	V	R	G	D	H	R	D	N	..	**F**	**D**	G	*P*	G	G	**N**	L	A	H	**A**	F	**Q**	..	T	N	N	F	R	E
MMP1+act	**F**	V	R	G	D	H	R	D	N	..	**F**	**D**	G	*P*	G	G	**N**	L	**A**	H	A	F	**Q**	..	T	N	N	F	R	E
MMP1+SIMP1	**F**	V	*R*	**G**	**D**	H	R	D	**N**	..	**F**	**D**	G	*P*	G	**G**	**N**	*L*	*A*	*H*	**A**	*F*	*Q*	..	T	*N*	**N**	**F**	R	**E**
MMP1+gal	F	V	R	G	D	H	R	D	**N**	..	F	**D**	G	*P*	**G**	G	*N*	*L*	*A*	**H**	A	F	Q	..	T	N	N	F	R	E

seq	210																				230		235							243	249
MMP12+NNGH	**T**	**N**	L	F	L	**T**	**A**	V	H	E	I	G	H	S	L	G	L	G	H	**S**	S	..	V	M	**F**	*P*	**T**	**Y**	K	**Y**	V	**S**
MMP1+NNGH	Y	N	L	H	R	**V**	A	**A**	H	**E**	L	G	H	S	L	G	**L**	**S**	**H**	**S**	T	..	L	M	Y	P	**S**	**Y**	T	F	**S**	A
MMP1+act	Y	N	L	H	R	**V**	A	**A**	**H**	E	**L**	G	H	**S**	L	G	L	**S**	H	S	T	..	**L**	**M**	Y	*P*	S	Y	T	F	S	**A**
MMP1+SIMP1	Y	N	L	H	R	**V**	A	**A**	**H**	E	**L**	G	**H**	**S**	L	G	L	**S**	**H**	S	T	..	**L**	**M**	**Y**	*P*	**S**	**Y**	*T*	*F*	S	A
MMP1+gal	Y	N	L	H	R	**V**	**A**	A	**H**	E	**L**	G	H	**S**	L	G	L	**S**	**H**	S	T	..	*L*	**M**	**Y**	*P*	S	**Y**	T	*F*	**S**	A

[a] Residue numbers refer to the MMP-1 sequence. Chemical shifts for residues in italics are not available.

coordination to the catalytic zinc. The ligands in the second cluster are oriented similarly to those in the first cluster, but the hydroxamic acid is coordinated to zinc only though the carboxylic oxygen. The structures in the fourth cluster show coordination of the sulfonate oxygen (SO) atoms to zinc. In all cases the *p*-methoxy-phenyl group sits in the S1′ hydrophobic pocket. The position of the *i*-butyl group changes in the four adducts. In the first and second clusters it prevents the formation of hydrogen bonding between the hydroxamic HN and alanine 182 oxygen, whereas the latter hydrogen bond is present in the third cluster.

NOE restraints were obtained in the following way. In the ^{15}N NOESY−HSQC spectrum, a cross-peak at chemical shift of 11.7 ppm is present in the N leucine181 plane (Figure 4). Such a shift is too high to be assigned to a protein signal, as no tryptophan residue is close to the active site. Therefore, it was assigned to the unique amide proton of NNGH. Aromatic protons of NNGH are close to N of residues glycine 221, histidine 222, alanine 216, and arginine 214 according to the structures calculated by Autodock. We have searched in the spectrum all the long range NOEs between aromatic protons and amide groups of these residues. New peaks in the spectrum of the adduct that cannot be due to intraprotein interactions actually appear in the aromatic region (Figure 4) and were assigned as reported in Table 2.

The structural families obtained with Xplor-NIH starting from the first three lowest Autodock docking energy structures converged to the same conformation (see Figure 5). This conformation was similar to the conformation of the third Autodock cluster, with the exception that the sulfur oxygen H-bonded to alanine 182 was the most external oxygen atom rather than the internal one. The lowest total energies were −622, −617, and −615 kcal mol^{-1}, respectively. The lowest total energy of the structural family obtained with Xplor-NIH starting from the fourth Autodock structure was −592 kcal mol^{-1}. This adduct, slightly different from the other three for the fact that zinc coordination by hydroxamate was loose, can be excluded due to its larger energy.

No appreciable changes in the protein side chain positions are observed and thus further Autodock/Xplor-NIH cycles were not needed. Therefore, the structural family shown in Figure 3C represents an experimentally validated and unique structural model for the MMP-1−NNGH adduct.

MMP-1−Actinonin. Actinonin, whose molecular structure is reported in Chart 1, is a well-known inhibitor of aminopeptidases and peptide deformylase.[25] It is also a strong inhibitor for some MMPs, with a K_i of 300 nM for its adduct with MMP-1.[26]

Chemical shift perturbations again allow us to map the region of interest on the protein surface. Residues 215, 217, 218, 220, 223, and 227 forming the metal binding site, residues 235, 236, and 249 on the loop that covers the S1′ pocket, and residues 180 and 182 on the spatially close strand (see Figure 1D) define the ligand binding region and were used for the definition of the Autodock grid. Despite the incomplete correspondence of the affected residues with those found for the NNGH adduct, the resulting grid was quite similar. Four clusters were then calculated (docked conformations were again clustered according to a maximal RMSD of 1.0 Å, see Figure 6). In all structures the hydroxamate is bound to the catalytic zinc. However, whereas in the first two structures the pentyl group is located inside the S1′ hydrophobic pocket and the external propyl group is differently oriented, in the third and fourth structures the two groups are interchanged. The lowest docking energy for the structures in the four clusters were −19.91, −19.11, −18.82, and −18.62 kcal mol^{-1}, respectively.

Cross-peaks of all protons belonging to the ligand with the HN protein protons expected at distances shorter than 5 Å for one or another cluster were looked for in the ^{15}N NOESY−HSQC spectrum. Since the NH of tyrosine 240 has two unassigned cross-peaks at frequencies typical of methyls, they must be related to two methyl groups that are close in the structure of the adduct. From the clusters generated by Autodock, they can only be H′ and H‴. The following peaks were thus assigned: (a) methyl protons H‴ with tyrosine 240 and with the aligned threonine 241 and (b) H′ with tyrosine 240. Among the clusters generated by Autodock, the third and fourth clusters can be readily excluded, because in such structures the above cross-peaks could not be observed. Therefore, by looking at the other two clusters, we also assigned the following cross-peaks, which cannot be assigned to other intraresidue protons or to side chain protons of close residues: (c) alanine 184 with H1 and tyrosine 240 with H6, as such protons are the closest to the coupled HN protons; (d) leucine 181 and tyrosine 240 with H′, as they are aligned and close to one another.

Xplor-NIH calculations were thus performed to refine the selected Autodock structures. Actually, we per-

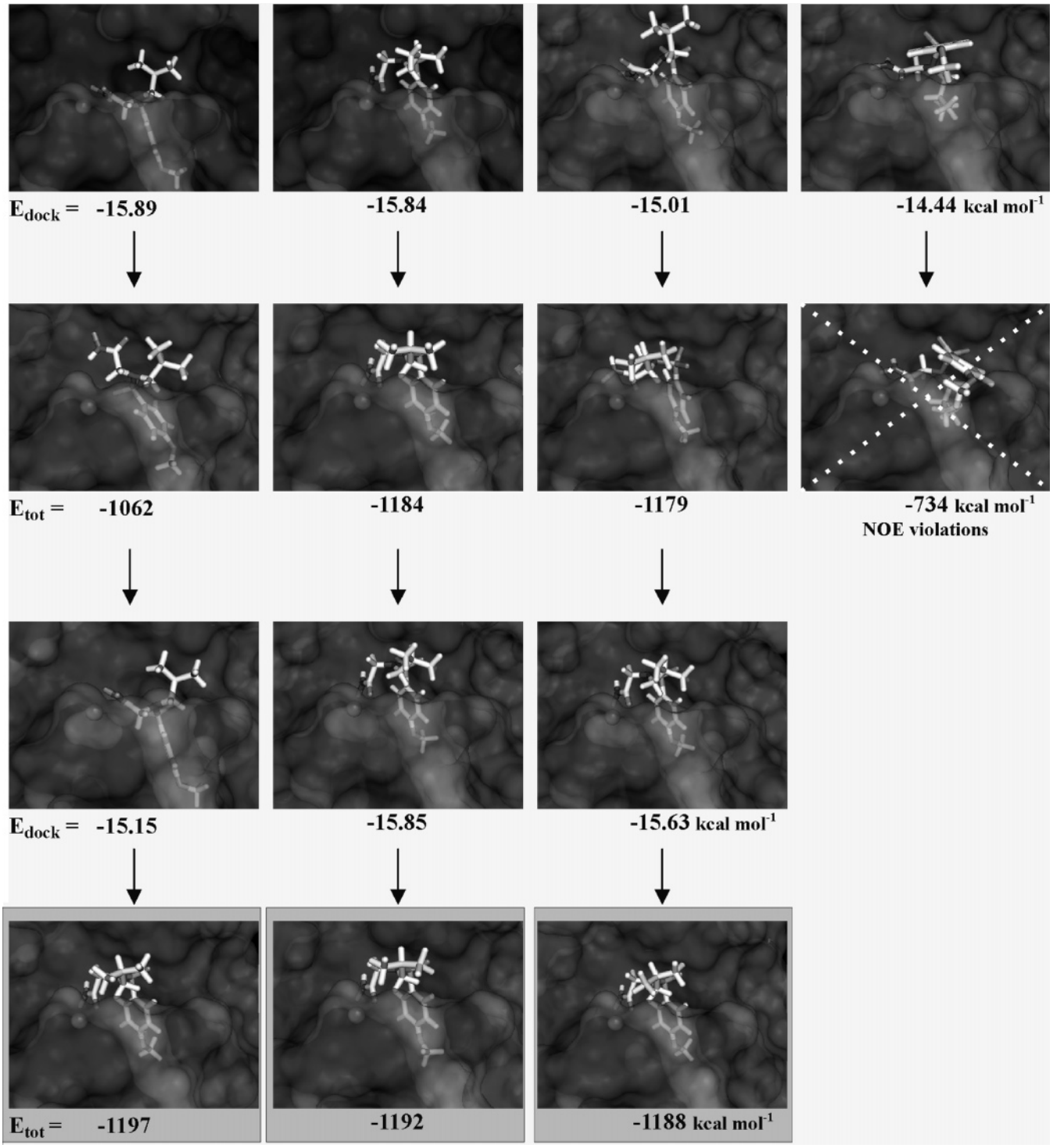


Figure 2. Representative structures of the MMP-12−NNGH adduct for the four lowest energy clusters obtained from Autodock (first row), Xplor-NIH (second row), a second Autodock run (third row), and further Xplor-NIH calculations (fourth row). The final validated structures are highlighted.

formed the calculations not only starting from the first two structures, but also starting from the structures excluded according to the observation of the ^{15}N NOESY−HSQC spectrum. The first and second family of structures calculated with Xplor-NIH are very similar to the corresponding Autodock structures; the third Xplor-NIH structural family shows significant rearrangements in the position of the ligand branches, but the pentyl group remains located outside the hydrophobic pocket; in the fourth Xplor-NIH structural family the pentyl group lies in the hydrophobic pocket, thus resulting similar to the first and second families. Xplor-NIH energies for the four families are −870, −862, −811, and −840 kcal mol^{-1}, respectively. This indicates that the third structure, quite different from the other three, is not acceptable. The calculations show that the method is indeed robust. In fact, the first two Autodock structures that were selected from the observation of the NMR spectra actually have the lowest energy, whereas the third has a sizably larger energy even after

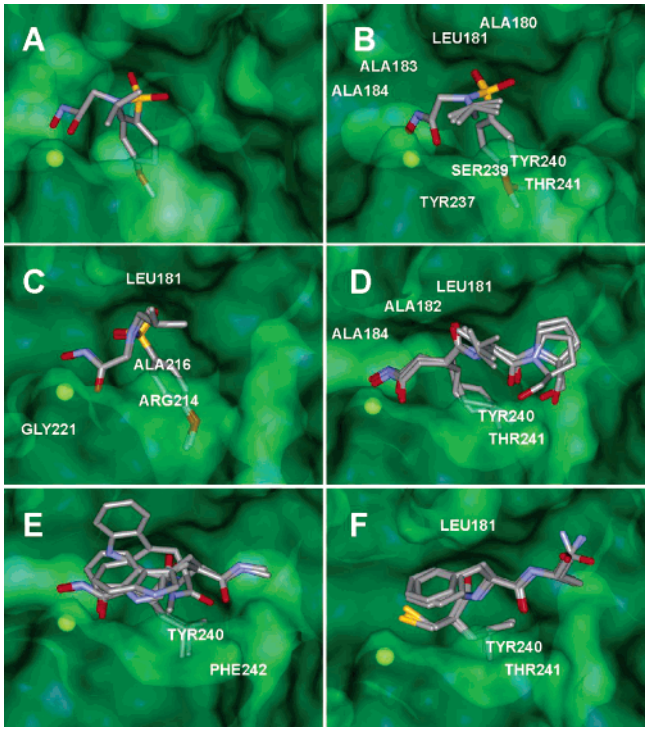


Figure 3. X-ray structure of the MMP-12−NNGH adduct (A), structures calculated with the proposed protocol of the MMP-12−NNGH adduct (B), and structures calculated with the proposed protocol for the adduct of MMP-1 with NNGH (C), actinonin (D), galardin (E), SIMP-1 (F). Labels in panels B−F indicate the residue numbers of amino acids exhibiting NOE contacts to the ligands.

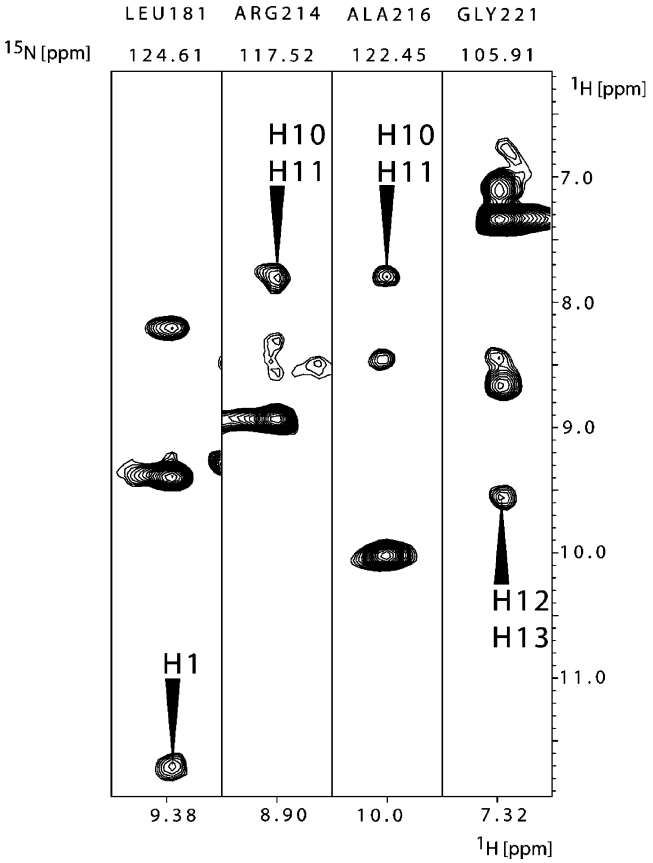


Figure 4. Protein−ligand cross-peaks observed in the ^{15}N NOESY−HSQC spectrum of the MMP-1−NNGH sample.

Xplor-NIH refinement. Interestingly, the fourth Autodock structure, initially completely different from the first two, was brought by Xplor-NIH calculations to converge with the first two.

A second Autodock and Xplor-NIH cycle was performed starting from the lowest energy protein structure. The calculated Xplor-NIH structures, in fact, showed slightly different positions of protein side chains, in particular of residues leucine 181, proline 238, and tyrosine 240. Such new protein conformation was provided to Autodock for a new docking calculation. The best four Autodock clusters (docking energy −20.61, −19.07, −18.99, and −18.88 kcal mol^{-1}) were then provided to Xplor-NIH. The first, third and fourth clusters display both the hydroxamate and the pentyl group in similar positions; the second one is completely different (the hydroxamate does not bind the zinc ion). All lowest energy Xplor-NIH structures (see Figure 3D), with the exception of those calculated starting from the second Autodock structures, converged to the third Autodock conformation, and are equivalent to the lowest energy Xplor-NIH family calculated in the first cycle. The total energies for these structures are −881, −880, and −870 kcal mol^{-1}. The Xplor-NIH structure calculated starting from the second Autodock structure has a total energy of −810 kcal mol^{-1} and can thus be excluded. Therefore, the structure family of Figure 3D is a unique structural model for the MMP-1−actinonin adduct.

MMP-1−Galardin. Galardin (see Chart 1) is a broad spectrum peptidomimetic inhibitor of MMPs[16] with an IC_{50} of 1.5 nM for MMP-1.[27] Chemical shift perturbation involved residues 215, 216, 218, 220, 223, 227, and 228 at the metal binding site, 236, 237, 240, and 243 at the large loop covering the S1′ cavity, and 179 and 183 at the strand facing the S1′ cavity and the metal binding site (Figure 1E). These residues were used to define the Autodock grid, which again was found very similar to the previous ones. The four lowest docking energy clusters calculated by Autodock (−19.79, −19.51, −18.59, and −17.48 kcal mol^{-1}, respectively) showed the following features (see Figure 7). In the first, second and third cluster the *i*-butyl group enters the S1′ pocket, whereas in the fourth cluster it is outside. The structures in the first and second clusters are very similar, as they differ only for the orientation of the indole group, positioned outside the S1′ pocket. The structures in the third and fourth cluster are quite different from those in the first and second cluster, including the position of the indole group, which in any case remains outside the S1′ pocket.

In the ^{15}N NOESY−HSQC spectrum, two cross-peaks are present in the N phenylalanine 242 and tyrosine 240 planes. Such peaks are at chemical shifts typical of methyl groups and cannot be assigned to any intraresidue proton or proton of close residues. Since in galardin there are three methyl groups, two of them being close in the structure, the latter (H′ and H″) were assigned to these peaks. Another cross-peak is present that cannot be assigned to protein protons in the plane of tyrosine 240. This cross-peak falls into the aliphatic region, and therefore it could be provided by CH or CH_2 protons. Since such proton must be close to H′ and H″, which have also a cross-peak with tyrosine 240, it was assigned to H6 or H7.

Table 2. Observed NOEs between MMP-1 Amide Protons and Ligand Protons

	181	182	184	214	216	221	240	241	242
NNGH	H1			(H10,H11)	(H10,H11)	H12 H13			
actinonin	H′	(H3,H4) H6	H1				H‴ H′ H6	H‴ (H23,H24)	
galardin							H′ H″ (H6,H7)		H′ H″
SIMP-1	(H17,H18,H19,H20,H21)						H′ H″	H″	

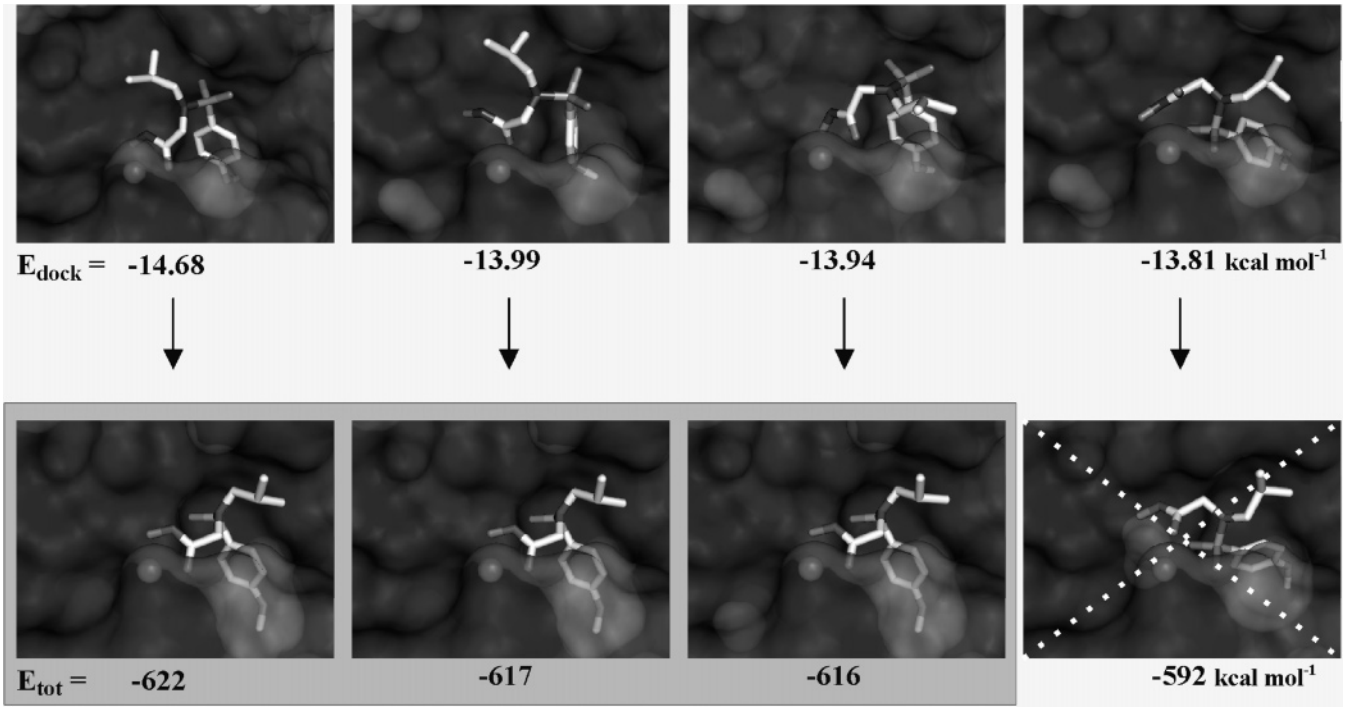


Figure 5. Representative structures of the MMP-1−NNGH adduct for the four lowest energy clusters obtained from Autodock (first row) and from Xplor-NIH (second row). At this point convergence was obtained. The final validated structures are highlighted.

Xplor-NIH calculations change only slightly the conformations obtained with Autodock relative to the first three clusters. The structure obtained starting from the fourth Autodock cluster is instead modified by the NOE restraints to have the *i*-butyl group inside the S1′ pocket as in the other three clusters. The total energy of the Xplor-NIH structures are −575, −578, −565, and −527 kcal mol^{-1}, respectively. Only small changes in the side chain positions have been observed, regarding in particular residues from 238 to 241.

The three protein structures with the smallest Xplor-NIH energy were used to repeat Autodock calculations. In the first case Autodock produced the two lowest docking energy clusters very similar to those obtained in the first run (−18.87 and −18.57 kcal mol^{-1}), whereas the third and fourth clusters (with docking energy −18.08 kcal mol^{-1}) have now the indole group inside the S1′ pocket. These conformations can be excluded by the observed NOEs. It is remarkable that such faulty Autodock behavior occurs in the second round, i.e., after adjustment of the structure by Xplor-NIH minimization. This observation underlines the need for experimental restraints to gain confidence in in silico models. In the second case, the three lowest docking energy clusters are again very similar to those obtained in the first run (−19.59, −19.56, and −18.14 kcal mol^{-1}), whereas in the fourth cluster (with docking energy −17.83 kcal mol^{-1}) the *i*-butyl group is outside the S1′ pocket. In the third case, the lowest docking energy cluster is again similar, with energy −18.42 kcal mol^{-1}. The Xplor-NIH calculations performed with the four lowest docking energy structures as starting conformations converged to a unique conformation (−586, −584, −576, and −576 kcal mol^{-1}), except for the indole group, which, being outside the S1′ pocket, is free to move (Figure 3E). Again, the family of Figure 3E can be confidently taken as a validated structural model for the galardin adduct of MMP-1.

MMP-1−SIMP-1. SIMP-1 is a polypeptide derivative able to inhibit collagenases.[16] Its molecular structure is reported in Chart 1. We measured an IC_{50} value for the adduct of 46 nM. In the MMP-1−SIMP-1 adduct, affected resonances include residues 215, 217−218, 220, 222, 223, 227, and 228 on the α helix of the zinc binding site, residues 235−237, 239, 240, and 242 forming the S1′ cavity, and residues 179, 180, and 184 on the strand facing both the catalytic metal and the S1′ pocket (Figure 1F). The four clusters with smallest docking energy calculated by Autodock (see Figure 8) have docking energy of −16.15, −15.92, −15.76, and −15.72

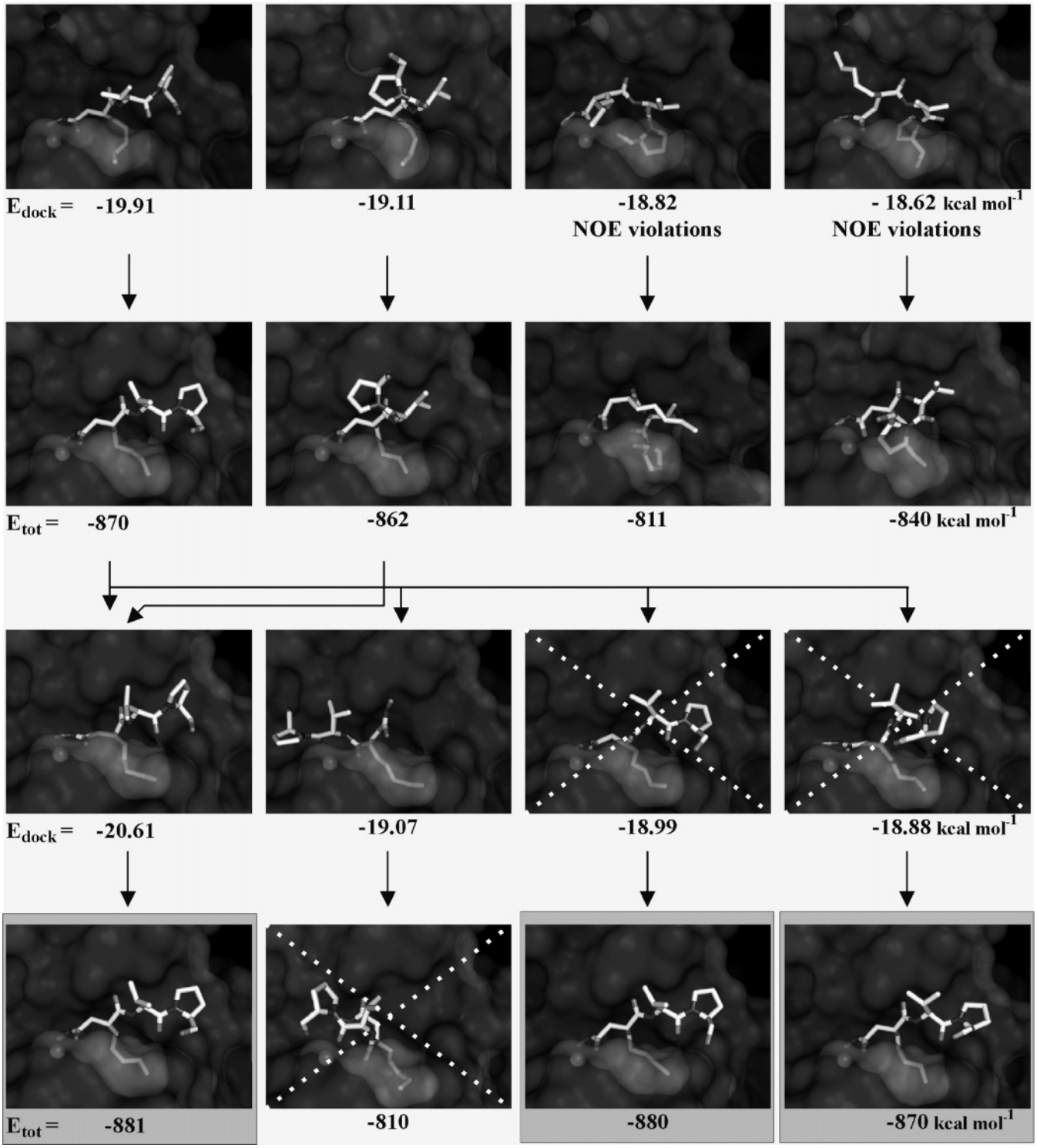


Figure 6. Representative structures of the MMP-1−actinonin adduct for the four lowest energy clusters obtained from Autodock (first row), Xplor-NIH calculations (second row), a second Autodock run (third row), and further Xplor-NIH calculations (fourth row). The final validated structures are highlighted.

kcal mol^{-1}. In the first and second clusters the sulfur atom coordinates the catalytic zinc; in the first cluster the S1′ pocket interacts with the ligand benzyl group, in the second with the *i*-butyl group. In the third cluster the sulfur atom is hydrogen bonded to the oxygen of glycine 179, on the other site of the catalytic pocket with respect to the zinc ion, and the ligand benzyl group sits in the S1′ pocket. In the fourth cluster, the ligand is oriented similarly as in the first cluster, but the ligand sulfur atom is loosely coordinated to the zinc ion, and hydrogen bonded to glutamate 219.

In the ^{15}N NOESY−HSQC spectrum, in the N plane of residue leucine 181, there are two signals in the aromatic region that cannot be assigned to protein side chains. Therefore, they must be assigned to protons of the aromatic ring of the SIMP-1. Two cross-peaks, one of low and one of high intensity, are present in the N plane of tyrosine 240 at frequencies typical of methyl

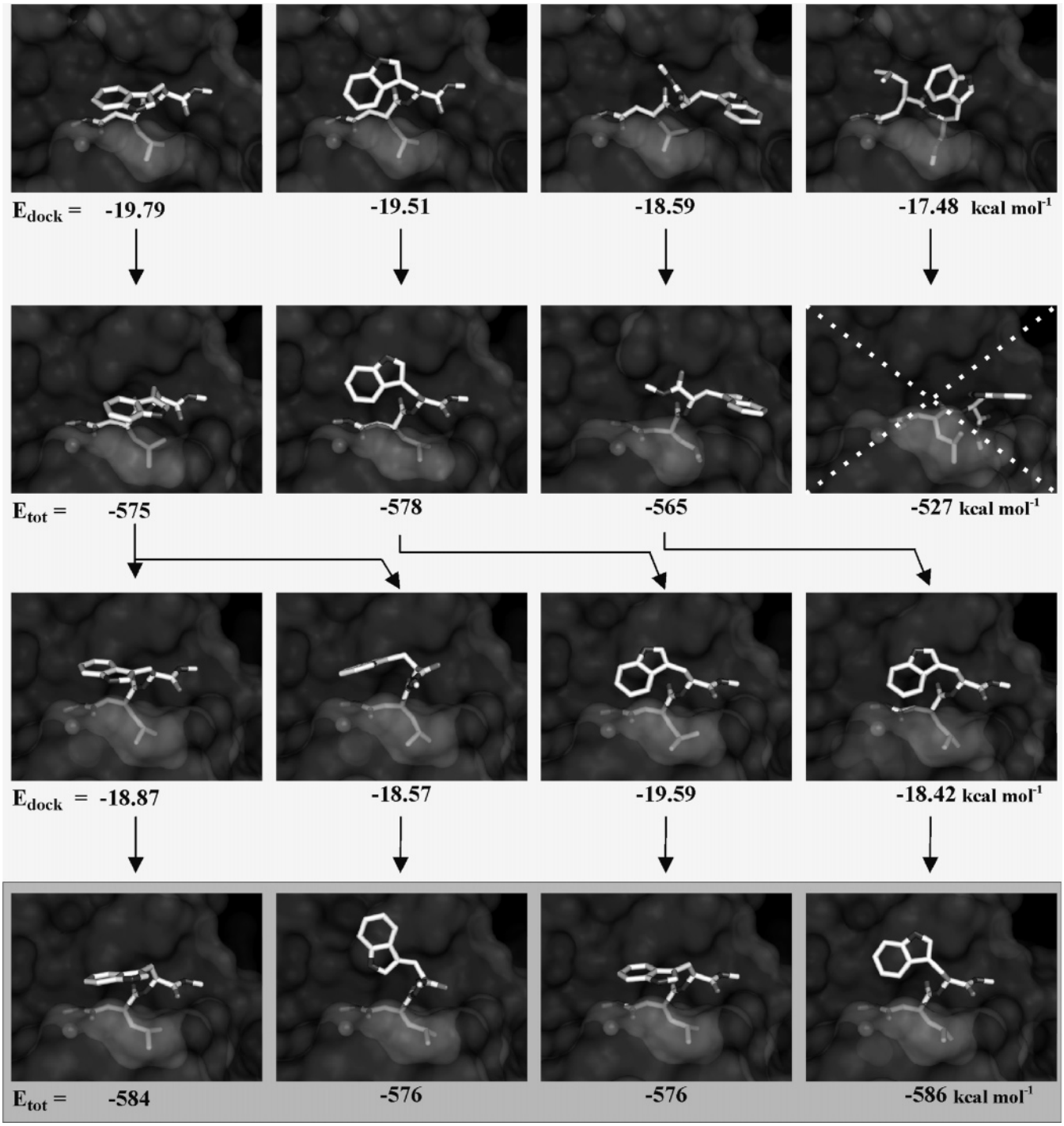


Figure 7. Representative structures of the MMP-1−galardin adduct for the four lowest energy clusters obtained from Autodock (first row), Xplor-NIH calculations (second row), a second Autodock run (third row), and further Xplor-NIH calculations (fourth row). The final validated structures are highlighted. The configuration of the indole ring is not defined because of lack of experimental restraints and strong energetic preference in Xplor-NIH calculations.

groups that cannot be assigned to intraresidue or sequential contacts. In one of the clusters calculated by Autodock, N of tyrosine 240 is close to two of the three methyls of SIMP-1, H′ being closest than H″, and thus the cross-peaks were correspondingly assigned (see Table 2). A further distance restraint is determined from another cross-peak in the N plane of threonine 241, aligned with the signal assigned to H″.

Xplor-NIH calculations select the second Autodock cluster as the correct one. In fact it remains almost unchanged after refinement, with total energy −709 kcal mol^{-1}. Calculations performed starting from the other clusters provide structures very different from the starting ligand conformation, and with the ligand not coordinated to the zinc ion. Their total energies are larger than −634 kcal mol^{-1} and such structures are thus excluded.

Slight changes in the protein side chain positions are observed, in particular on residues 180, 214 and 219. A second Autodock calculation was thus performed. The first three clusters (with docking energy of −16.15, −16.12 and −15.88 kcal mol^{-1}, respectively) show a

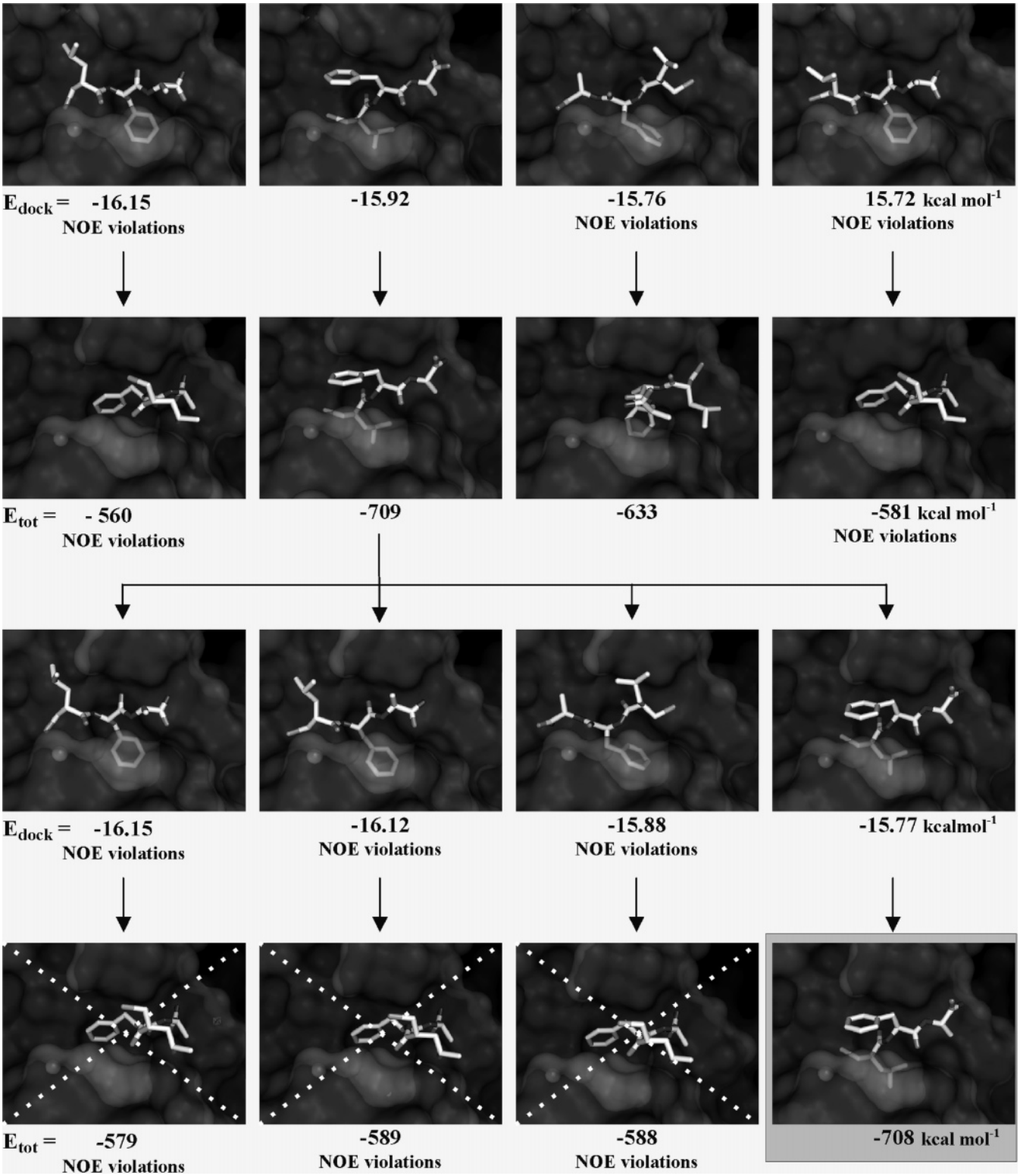


Figure 8. Representative structures of the MMP-1−SIMP-1 adduct for the four lowest energy clusters obtained from Autodock (first row), Xplor-NIH calculations (second row), a second Autodock run (third row), and further Xplor-NIH calculations (fourth row). The final validated structures are highlighted.

ligand pose similar to that calculated in the first and fourth clusters of the first Autodock run. The fourth cluster, with docking energy −15.77 kcal mol^{-1}, is instead similar to the pose already identified as correct. Xplor-NIH calculations again confirmed such structure as the correct one, with total energy −708 kcal mol^{-1}. The corresponding family is shown in Figure 3F. This family represents the validated structural model of the MMP-1−SIMP-1 adduct.

Backbone Mobility. To test the protocol for possible protein backbone rearrangements upon complexation, Xplor-NIH calculations were also performed with allowing the protein backbone to move in the protein region affected by chemical shift perturbation. In all

cases we found no appreciable differences in the results. In fact, for all adducts the lowest energy structures corresponded to those identified as correct in the calculations performed with rigid backbones.

Discussion

A protocol has been developed to merge the "pure" docking capability of Autodock (or other docking programs) with the exploitation of available experimental restraints. For the relatively strong ligands (K_{diss} approximately micromolar or less) elected here, the protocol has been shown to be efficient, robust and reliable. As shown in Scheme 1, the protein binding site is identified from chemical shift perturbation in the HSQC spectrum of the protein upon complexation. The observation of shift perturbations on passing from the assigned spectrum of the free protein to the spectrum of the adduct permits the definition of the protein grid to be used in Autodock calculations. Autodock usually provides several clusters of structures for the adduct, which often have similar docking energy. These structures are used to calculate maps of NOEs, to be compared with NOEs actually observed in the ^{15}N NOESY−HSQC spectrum of the adduct. A few ligand−protein NOEs can always be assigned, and the latter can be used as restraints in Xplor-NIH calculations for selection, validation and refinement of the Autodock structures. One-two cycles at most may be needed in case Xplor-NIH calculations modify some protein side chain positions with respect to the structure provided to Autodock. All these steps could be performed semiautomatically, if required.

The protocol relies on the following information to be available: the protein structure; the assigned HSQC spectrum of the free protein; the ^{15}N NOESY−HSQC spectrum of the free protein; the HSQC and ^{15}N NOESY−HSQC spectra of the protein−ligand adduct; and the 1D ^{1}H spectrum of the ligand. The protocol has been developed in order to avoid preparation of doubly labelled samples and assignment of protein side chains, thus resulting in a much faster throughput.

We have shown that such approach is actually efficient in finding the protein−ligand structure for four adducts of MMP-1 with different ligands. The peculiarity that makes this approach successful in the cases here examined is the combination of a docking program, able to quickly and efficiently sample the possible binding poses, with a molecular dynamics program, which selects the proposed poses using few unambiguous experimental data. In this way the efficiency of the former program is coupled to the complexity of the latter, which also allows for protein side chain movements. The program has been deliberately tested using only unambiguous NOEs obtainable from the assignment of HN, but it is obviously open to the use of additional or different restraints. We decided to use the chemical shift perturbations only for the determination of the grid to be used for the docking program calculations, without including them as restraints in the molecular dynamics program due to their ambiguous nature, although ambiguous restrains could be in principle used, either as such, as recently proposed,[13] or through calculation of *j*-surfaces.[10] The use of chemical shift perturbations for the determination of the grid is much less stringent than their use as constraints, as a few "second sphere" shifts erroneously mistaken for first sphere shifts may drive the ligand in wrong positions, while the resulting grids are expected to be only somewhat broadened. As a matter of fact, differences in perturbed residues from one ligand to another do not result in grossly different grids, and the latter, in all cases, encompassed the whole catalytic site.

Several predicting programs for protein−ligand adducts have been proposed in the literature. Inclusion of biochemical and biophysical data in docking protocols, called guided docking,[28,29] is a common approach to reduce the conformational variety of the proposed solutions. Some other programs[7−9,30−35] work totally in silico, without experimental information on the investigated adduct, and perform docking calculations with an improved level of sophistication. They can be successful, but the level of confidence for the proposed adduct is difficult to establish. Furthermore, a strong bias toward known solutions or preconceived requirements is introduced if the docking is restrained according to chemical information derived from databases of protein−ligand complexes. Other programs[36−39] use the experimental NMR information more systematically, thus being similar to structural determination programs and therefore more time-consuming. NMR-derived restraints were also used in docking programs to identify the location of the ligand binding[10] and to restrict the conformational space for molecular modeling routines.[11] NMR experiments on selectively labeled proteins were also used to obtain structural information on protein−ligand complexes.[12] This approach, although more expensive than the one here proposed, is probably the only viable in case of large proteins. To our knowledge this is the first time that an approach is proposed where few experimental data are used to select and refine poses proposed by fast docking programs.

Autodock has been selected among the docking programs because in the case of MMPs it was demonstrated to be a robust program with good docking accuracy and reliability, including the correct geometry of the zinc binding groups.[21,40] It employs a genetic algorithm searching function, able to efficiently sample large search spaces. Different docking programs could however be used if considered more reliable in other cases. In the same way, other molecular dynamic programs could be used instead of Xplor-NIH. We used Xplor-NIH as an NMR-oriented widespread general program for structural calculations using simulated annealing. Ligand growing procedures[30] may also be implemented in Xplor-NIH, resulting probably useful especially in case of large ligands.

Although the presence of the metal ion in MMPs tends to restrict the number of Autodock clusters by favoring poses where the hydroxamic moiety is coordinated to the metal, the protocol is expected to be useful also in case of proteins not containing catalytic ions. Actually, docking programs are developed to work mainly in their absence, and, in case they propose several different conformations, the detection of NOEs may result decisive for the selection of the correct one. Indeed, as we have seen, Autodock does not always succeed in correctly binding the metal to the hydroxamic moiety. Furthermore, in the absence of the metal, further

H-bonds or van der Waals contacts should occur for strong ligands, which would likely provide additional intermolecular NOEs.

We have shown that it is possible to obtain few intermolecular experimental NOEs through fast NMR experiments without the necessity to assign all protein NOESY cross-peaks. Only unambiguous NOEs between protein and ligand protons have been considered; therefore, cross-peaks were assigned to ligand protons only if they could not be reasonably assigned to any protein side chain proton, taking into account the structural adducts proposed by Autodock. In all cases here addressed, experimental restraints have been shown to be necessary and sufficient to extract the adduct conformation among the several proposed by Autodock with similar docking energy, and thus are used to validate them. Furthermore, the approach proposed can also be useful to refine the structure of the ligand−protein adduct, especially because local small modifications in the protein structure (of side chains, if sufficient as in the present case, but also in the protein backbone, if needed − see below) can be accommodated by cycling between Autodock/Xplor-NIH runs. This makes the present approach preferable to the direct introduction of distance restraints in docking programs with a fixed protein matrix.

The solution structure of the inhibitor-free MMP-1, obtained from a series of 3D triple-resonance NMR experiments, shows nearly identical both backbone and secondary structures than the crystallographic structures.[41] Furthermore, the backbones of the solution structures of the inhibitor-free MMP-1 and of the MMP-1 complexed with a sulfonamide derivative of the hydroxamic acid compound have been shown to be essentially identical,[42] although mobility measurements indicate that the region near the active site is highly mobile.[41,42] It is thus reasonable, at least in our case, to assume that the protein backbone remains rigid during complexation in solution, and with structure identical to the crystallographic structure.

Although not necessary for the present calculations, also the protein backbone could be allowed to (partially) move in Xplor-NIH calculations (see results). This could be important if modest backbone rearrangements are expected upon ligand binding, as could be indicated by chemical shift perturbations spread out over a wider region.

It is known that effective MMP inhibitors achieve tight binding via extensive van der Waals contacts with the hydrophobic interior of S1′ and by strong electrostatic interactions with zinc and nearby charged or polar side chains.[43] All calculated adducts indeed show ligand coordination to the catalytic zinc and the formation of a net of hydrogen bonds between ligand and protein residues. This result is not trivial as it may seem, as several of the initially obtained Autodock structures had severely distorted − or were even lacking − hydroxamate coordination to the zinc ion.

The distance between zinc and hydroxamate oxygens is in all calculated structures between 1.95 and 2.25 Å. The O−Zn−O angle is always between 86 and 93°. The coordination geometry is distorted square-pyramidal in MMP-12−NNGH and MMP-1−actinonin, and distorted trigonal bipyramidal, with hydroxamic O2 and N histidine 222 in axial positions, in MMP-1−NNGH and galardin. All hydrogen bonding interactions between MMPs and ligands are reported in Table 3. In particular, H-bonds are present in all adducts with NNGH, actinonin, and galardin between oxygen of alanine 182 and the amide proton of the hydroxamic group, as well as between the protonated glutamate 219 and the oxygen of the hydroxamic group. H bonds are also present between ligands and HN of Leu 181, as previously seen in the MMP-1-CGS[42] and in the MMP-12−NNGH adducts. In the MMP-1−SIMP-1 adduct, with a distorted tetrahedral coordination geometry around the zinc ion, constituted by the three histidine nitrogen atoms and the sulfur SIMP-1 atom, a net of hydrogen bonds is formed, connecting the ligand to the protein atoms (see Table 3). Both the coordination geometry and the H-bonding network can be used to assess the reliability of the obtained adducts. In all the adducts, the inhibitors establish enough interactions to reach nanomolar affinity. In particular, all the ligands bind the metal, place a lipophilic moiety into the S1′ cavity and establish two or more hydrogen bonds with atoms of the protein groove. This binding mode is reasonable and is indeed adopted by many strong ligands of MMPs.

Table 3. Predicted H-bonds between Protein and Ligand Nuclei

MMP-12−NNGH	181 Leu HN	NNGH O4/O3
	182 Ala HN	NNGH O3
	219 Glu HO1	NNGH O1
	NNGH H1	182 ALA O
MMP-1−NNGH	182 Ala HN	NNGH O4
	219 Glu HO1	NNGH O1
	NNGH H1	182 Ala O
MMP-1−actinonin	181 Leu HN	act O3
	219 Glu HO1	act O1
	240 Tyr HN	act O4
	act H1	182 Ala O
	act N2	238 Pro O
MMP-1−galardin	181 Leu HN	gal O3
	182 Ala HN	gal O3
	219 Glu HO1	gal O1
	240 Tyr HN	gal O4
	gal H5	238 Pro O
	gal H1	182 Ala O
MMP-1−SIMP-1	181 Leu HN	SIMP1 O1
	182 Ala HN	SIMP1 O1
	240 Tyr HN	SIMP1 O2
	SIMP1 H8	238 Pro O
	SIMP1 H1	219 Glu OE1
	SIMP1 H23	179 Gly O
	SIMP1 H26	210 Tyr OH

Cycling between fast docking programs and Xplor-NIH calculations can be used to assess ligand−protein structures also in the presence of restraints different from NOEs. Diamagnetic residual dipolar couplings have already been demonstrated to be extremely useful to predict the structure of protein−protein adducts.[20,44−46] Also pseudocontact shifts have been used for the study of protein−protein docking.[47] Paramagnetism-based restraints, and in particular paramagnetic relaxation rates, pseudocontact shifts and residual dipolar couplings, arising when a paramagnetic metal ion is coordinated to the protein, could be employed as restraints in the proposed protocol for protein−ligand docking. Xplor-NIH has the advantage that it already contains the tools needed to deal with such restraints.[48]

Conclusions

A novel protocol to obtain validated structural models of protein−ligand complexes has been developed and applied for the determination of the structure of the adducts of the protein MMP-1 with four different ligands. The method was shown to be reliable, as tested for the known structure of the adduct of one of these ligands with MMP-12. It uses NMR derived restraints obtained using singly (^{15}N) labeled proteins. The strategy that we propose promises to be generally useful also for the structural determination of different protein−ligand adducts, whenever the structure of the free protein is known and the structural changes upon complexation are not expected to be dramatic.

Experimental Section

Sample Preparation. The fragment of human fibroblast collagenase corresponding to proMMP-1 (Pro21-Pro269) and bearing an additional methionine at the N-terminal, was expressed in *Escherichia coli*. The cDNA was cloned into the pET21 vector (Novagen) using *Nde*I and *Xho*I as restriction enzymes. The *E. coli* strain BL21 Codon Plus cells, transfected with the above vector, were grown in 2 × YT media at 37 °C. The protein expression was induced during the exponential growth phase with 0.5 mM of IPTG. Cells were harvested for 4 h after induction. Uniform^{15}N-labeled protein was obtained by growing the transfected BL21 Codon Plus cells in minimal media at 37 °C. The cells were lysed by sonication and the inclusion bodies, containing the proMMP-1, were solubilized in 2 M urea, 20 mM Tris (pH 8.0). The protein was purified on the Hitrap Q column (Pharmacia) with a buffer containing 2 M urea and 20 mM Tris (pH 8.0). The elution was performed using a linear gradient of NaCl up to 0.35 M. The purified protein was then refolded by using a multistep dialysis against solutions containing 50 mM Tris (pH 7.2), 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl. The refolded protein was exchanged, by dialysis, against a buffer with 10 mM Tris (pH 7.2), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl. The protein was activated by 1 mM APMA (4-aminophenylmercuric acetate) at 4 °C overnight and dialyzed with a buffer containing 10 mM Tris (pH 7.2), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl, 0.2 M acetohydroxamic acid (AHA). The activated protein (Val 101−Pro 269) was concentrated using an Amicon stirrer and Centriprep concentrators, fitted with a YM10 membrane in nitrogen atmosphere at 4 °C. Catalytic domain of MMP-1 was purified using size-exclusion chromatography with the final dialysis buffer and concentrated up to 0.5 mM using an Centriprep concentrators in nitrogen atmosphere at 4 °C. The final protein sample was dialyzed against a solution containing 50 mM sodium acetate, 100 mM NaCl, 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, with 10% of D_2O (pH 6.5).

Inhibited proteins were prepared by titration of the free-MMP-1 with equimolar amounts of NNGH, SIMP-1, galardin, and actinonin.

NNGH, galardin, and actinonin were purchased by BIOMOL International; SIMP-1 was purchased by Peptide International, Inc.

In Vitro Assay. The compounds were evaluated for their ability to inhibit the hydrolysis of fluorescence-quenched peptide substrate Mca-Pro-Leu-Gly-Leu-Dpa-Ala-Arg-NH_2 (Biomol, Inc.). The assays were performed in 50 mM HEPES buffer, containing 10 mM $CaCl_2$, 0.05% Brij-35, at pH 7, using 1 nM of MMP-1 catalytic domain and 1 μM of peptide. The enzyme was incubated at 25 °C with increasing concentration of inhibitor and the florescence ($excitation_{max}$ 328 nm; $emission_{max}$ 393 nm) was measured for 3 min after the addition of the substrate using a Varian Eclipse fluorimeter. Fitting of rates as a function of inhibitor concentration provided the IC_{50} values. In our experimental conditions with low enzyme concentration and peptide concentration much lower than K_M (the concentration of the substrate that leads to half-maximal velocity of the enzymatic hydrolysis reaction); the IC_{50} values provide a good estimate of the dissociation constant of the adduct. The inhibitor *N*-isobutyl-*N*-[4-methoxyphenylsulfonyl]-glycyl hydroxamic acid (Biomol, Inc.) was used as control.

NMR Measurements. ^{1}H ^{15}N HSQC experiments implemented with the sensitivity enhancement scheme[49] and ^{15}N NOESY−HSQC spectra[50] were performed on the free MMP-1 catalytic domain and on each protein−ligand adduct. ^{15}N NOESY−HSQC experiments were acquired with a mixing time of 110 ms and with data sets comprising 256(^{1}H) × 64(^{15}N) × 2048(^{1}H) data points. The NMR spectra were recorded on Avance 900 Bruker spectrometer, operating at proton nominal frequencies of 900.13 MHz and equipped with a triple resonance cryoprobe. All NMR experiments, recorded at 298 K, were processed using the standard Bruker software (XWIN-NMR), and analyzed through the XEASY program.[51]

Computer Programs. Autodock 3.0.5 was used to predict protein−ligand docking. It uses a Lamarckian genetic algorithm as global optimizer combined with energy minimization as a local search method.[52] Its scoring function is provided by the sum, with empirically determined scaling factors, of a Lennard-Jones 12−6 dispersion/repulsion term, a directional 12−10 hydrogen bond term, a Coulombic electrostatic potential, a term related to unfavorable entropy due to restrictions in conformational degree of freedom of the ligand, and a desolvation term. The PDB file was processed by Autodock Tool Kit. Reliable zinc parameters were provided as in ref 21. A box of 70 × 70 × 70 points with a grid spacing of 0.375 Å was defined as docking space. The ligands were generated and minimized using semiempirical calculations (AM1 type Gaussian 98),[53] and the pdbq files, comprising all protons, were provided to Autodock after all the Gasteiger−Marseli charges[54] were assigned by BABEL. For each run, a maximum number of 28 000 genetic algorithm operations were generated on a single population of 50 individuals. For each ligand, a total of 100 docking runs were performed, and the results were ranked according to the docking energy. Crossover, mutation, and elitism weights were set to 0.80, 0.02, and 1, respectively.

All minimization and dynamics calculations were carried out using the program Xplor-NIH.[55,56] The parameter and topology files for the ligands were generated using Xplo2D,[57] the improper angles being manually edited and the dihedral angles being set with force constant equal to zero. Protein electrostatic and van der Waal energy parameters have been evaluated using CHARMM nonbonded parameters.[58]

Acknowledgment. This work was supported by the European Union (Contract QLG2-CT-2002-00988 and Contract LSHG-CT-2004-512077), by MIUR−FIRB RBNE01TTJW, by MIUR−FISR "Modeling di strutture di metalloproteine e delle interazioni proteina-farmaco e proteina-proteina", and by Ente Cassa di Risparmio di Firenze.

References

(1) Anderson A. C. The Process of Structure-Based Drug Design. *Chem. Biol.* **2003**, *10*, 787−797.
(2) Blundell, T. L.; Jhoti, H.; Abell, C. High-throughput crystallography for lead discovery in drug design. *Nat. Rev. Drug Discovery* **2002**, *1*, 45−54.
(3) Lepre, C. A.; Moore, J. M.; Peng, J. W. Theory and Applications of NMR−Based Screening in Pharmaceutical Research. *Chem. Rev.* **2004**, *104*, 3641−3675.
(4) Pellecchia, M.; Becattini, B.; Crowell, K. J.; Fattorusso, R.; Forino, M.; Fragai, M.; Jung, D.; Mustelin, T.; Tautz L. NMR-based techniques in the hit identification and optimisation processes. *Exp. Opin. Ther. Targets* **2004**, *6*, 597−611.
(5) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **1996**, *274*, 1531−1534.
(6) Bertini, I.; Calderone, V.; Cosenza, M.; Fragai, M.; Lee, Y.-M.; Luchinat, C.; Mangani, S.; Terni, B.; Turano, P. Conformational variability of MMPs: beyond a single 3D structure. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5334−5339.

(7) Stahl, M.; Rarey, M. Detailed Analysis of Scoring Functions for Virtual Screening. *J. Med. Chem.* **2001**, *44*, 1035−1042.
(8) Cummings, M. D.; DesJarlais, R. L.; Gibbs, A. C.; Mohan, V.; Jaeger, E. P. Comparison of Automated Docking Programs as Virtual Screening Tools. *J. Med. Chem.* **2005**, *48*, 962−976.
(9) Trosset, J.-Y.; Scheraga, H. A. Reaching the global minimum in docking simulations: A Monte Carlo energy minimization approach using Bezier splines. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8011−8015.
(10) McCoy, M. A.; Wyss, D. F. Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. *J. Am. Chem. Soc.* **2002**, *124*, 11758−11763.
(11) Lugovskoy, A. A.; Degterev, A. I.; Fahmy, A. F.; Zhou, P.; Gross, J. D.; Yuan, J.; Wagner, G. A Novel Approach for Characterizing Protein Ligand Complexes: Molecular Basis for Specificity of Small-Molecule BcI-2 Inhibitors. *J. Am. Chem. Soc.* **2002**, *124*, 1234−1240.
(12) Pellecchia, M.; Meininger, D.; Dong, Q.; Chang, E.; Jack, R.; Sem, D. S. NMR-based structural characterization of large protein−ligand interactions. *J. Biomol. NMR* **2002**, *22*, 165−173.
(13) Schieborr, U.; Vogtherr, M.; Elshorst, B.; Betz, M.; Grimme, S.; Pescatore, B.; Langer, T.; Saxena, K.; Schwalbe, H. How much NMR-Data are required for the Determination of a Protein−Ligand Complex? *ChemBioChem* **2005**, *6*, 1891−1898.
(14) Woessner, J. F., Jr. Matrix metalloproteinases and their inhibitors in connective tissue remodeling. *FASEB J.* **1991**, *5*, 2145−2154.
(15) Yu, A. E.; Hewitt, R. E.; Connor, E. W.; Stetler-Stevenson, W. G. Matrix metalloproteinases: novel targets for directed cancer therapy. *Drugs Aging* **1997**, *11*, 229−244.
(16) Whittaker, M.; Floyd, C. D.; Brown, P.; Gearing, A. J. Design and Therapeutic Application of Matrix Metalloproteinase Inhibitors. *Chem. Rev.* **1999**, *99*, 2735−2776.
(17) Clendeninn, N. J.; Appelt, K. *Metalloproteinase Inhibitors in Cancer Therapy*; Humana Press: Totowa, NJ, 2001.
(18) Maeda, A.; Sobel, R. A. Matrix metalloproteinases in the normal human nervous system, microglial nodules, and multiple sclerosis lesions. *J. Neuropathol. Exp. Neurol.* **1996**, *55*, 300−309.
(19) Mitton-Fry, R. M.; Anderson, E. M.; Hughes, T. R.; Lundblad, V.; Wuttke, D. S. Conserved Structure for Single-Stranded Telomeric DNA Recognition. *Science* **2002**, *296*, 145−147.
(20) Clore, G. M.; Schwieters, C. D. Docking of Protein−Protein Complexes on the Basis of Highly Ambiguous Intermolecular Distance Restraints Derived from 1HN/15N Chemical Shift Mapping and Backbone 15N-1H Residual Dipolar Couplings Using Conjoined Rigid Body/Torsion Angle Dynamics. *J. Am. Chem. Soc.* **2003**, *125*, 2902−2912.
(21) Hu, X.; Shelver, W. H. Docking studies of matrix metalloproteinase inhibitors: zinc parameter optimization to improve the binding free energy prediction. *J. Mol. Graphics Modell.* **2003**, *22*, 115−126.
(22) MacPherson, L. J.; Bayburt, E. K.; Capparelli, M. P.; Carroll, B. J.; Goldstein, R.; Justice, M. R.; Zhu, L.; Hu, S.; Melton, R. A.; Fryer, L.; Goldberg, R. L.; Doughty, J. R.; Spirito, S.; Blancuzzi, V.; Wilson, D.; O'Byrne, E. M.; Ganu, V.; Parker, D. T. Discovery of CGS 27023A, a nonpeptidic, potent, and orally active stromelysin inhibitor that blocks cartilage degradation in rabbits. *J. Med. Chem*. **1997**, *40*, 2525−2532.
(23) Bertini, I.; Calderone, V.; Fragai, M.; Luchinat, C.; Mangani, S.; Terni, B. X-ray structures of ternary enzyme−product-inhibitor complexes of MMP. *Angew. Chem., Int. Ed.* **2003**, *42*, 2673−2676.
(24) Lovejoy, B.; Hassell, A. M.; Luther, M. A.; Weigl, D.; Jordan, S. R. Crystal Structures of Recombinant 19-kDa Human Fibroblast Collagenase Complexed to Itself. *Biochemistry* **1994**, *33*, 8207−8217.
(25) Chen, D. Z.; Patel, D. V.; Hackbarth, C. J.; Wang, W.; Dreyer, G.; Young, D. C.; Margolis, P. S.; Wu, C.; Ni, Z.-J.; Trias, J.; White, R. J.; Yuan, Z. Actiononin, a Naturally Occurring Antibacterial Agent, Is a Potent Deformylase Inhibitor. *Biochemistry* **2000**, *39*, 1256−1262.
(26) BIOMOL Research Laboratories data (www.biomol.com), 2005.
(27) Yamamoto, M.; Tsujishita, H.; Hori, N.; Ohishi, Y.; Inoue, S.; Ikeda, S.; Okada, Y. Inhibition of Membrane-Type 1 Matrix Matalloproteinase by Hydroxamate Inhibitors: An Examination of the subsite pocket. *J. Med. Chem.* **1998**, *41*, 1209−1217.
(28) van Dijk, A. D. J.; Boelens, R.; and Bonvin, A. M. J. J. Data-driven docking for the study of biomolecular complexes. *FEBS J.* **2005**, *272*, 293−312.
(29) Fradera, X.; Mestres, J. Guided Docking Approaches to Structure-Based Design and Screening. *Curr. Top. Med. Chem.* **2004**, *4*, 687−700.
(30) Ota, N.; Agard, D. A. Binding Mode Prediction for a Flexible Ligand in a Flexible Pocket using Multi-Conformation Simulated Annealing Pseudo Crystallographic Refinement. *J. Mol. Biol.* **2001**, *314*, 607−617.
(31) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. A review of protein-small molecule docking methods. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151−166.
(32) Perola, E.; Walters, W. P.; Charifson, P. S. A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* **2004**, *56*, 235−249.
(33) Kellenberger, E.; Rodrigo, J.; Muller, P.; Rognan, D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins* **2004**, *57*, 225−242.
(34) Carlson, H. A. Protein flexibility and drug design: how to hit a moving target. *Curr. Opin. Chem. Biol.* **2002**, *6*, 447−452.
(35) Gervasio, F. L.; Laio, A.; Parrinello, M. Flexible docking in solution using metadynamics. *J. Am. Chem. Soc.* **2005**, *127*, 2600−2607.
(36) Maurer, M. C.; Trosset, J. Y.; Lester, C. C.; DiBella, E. E.; Scheraga, H. A. New general approach for determining the solution structure of a ligand bound weakly to a receptor: structure of a fibrinogen Aα-like peptide bound to thrombin-(S195A) obtained using NOE distance constraints and an ECEPP/3 flexible docking program. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 29−48.
(37) Vogtherr, M.; Fiebig, K. NMR-based screening methods for lead discovery. *EXS* **2003**, *93*, 183−202.
(38) Zabell, A. P. Z.; Post, C. B. Docking multiple conformations of a flexible ligand into a protein binding site using NMR restraints. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 295−307.
(39) Hajduk, P. J.; Mack, J. C.; Olejniczak, E. T.; Park, C.; Dandliker, P. J.; Beutel, B. A. SOS−NMR: A saturation transfer NMR-based method for determining the structures of protein−ligand complexes. *J. Am. Chem. Soc.* **2004**, *126*, 2390−2398.
(40) Hu, X.; Balaz, S.; Shelver, W. H. A practical approach to docking of zinc metalloproteinase inhibitors. *J. Mol. Graphics Modell.* **2004**, *22*, 293−307.
(41) Moy, F. J.; Pisano, M. R.; Chanda, P. K.; Urbano, C.; Killar, L. M.; Sung, M. L.; Powers, R. Assesignments, secondary structure and dynamics of the inhibitor-free catalytic fragment of human fibroblast collagenase. *J. Biomol. NMR* **1997**, *10*, 9−19.
(42) Moy, F. J.; Chanda, P. K.; Chen, J. M.; Cosmi, S.; Edris, W.; Skotnicki, J. S.; Wilhelm, J.; Powers, R. NMR Solution Structure of the Catalytic Fragment of Human Fibroblast Collagenase Complexed with a Sulfonamide Derivative of a Hydroxamic Acid Compound. *Biochemistry* **1999**, *38*, 7085−7096.
(43) Wasserman, Z. R. Making a New Turn in Matrix Metalloprotease Inhibition. *Chem. Biol.* **2005**, *12*, 143−148.
(44) Clore, G. M. Accurate and rapid docking of protein−protein complexes on the basis of intermolecular nuclear overhauser enhancement data and dipolar couplings by rigid body minimization. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9021−9025.
(45) Clore, G. M.; Bewley, C. A. Using Conjoined Rigid Body/Torsion Angle Simulated Annealing to Determine the Relative Orientation of Covalently Linked Protein Domains from Dipolar Couplings. *J. Magn. Reson.* **2002**, *154*, 329−335.
(46) Jain, N. U.; Wyckoff, T. J. O.; Raetz, C. R. H.; Prestegard, J. H. Rapid Analysis of Large Protein−Protein Complexes Using NMR-derived Orientational Constraints: The 95 kDa Complex of LpxA with Acyl Carrier Protein. *J. Mol. Biol.* **2004**, *343*, 1379−1389.
(47) Crowley, P. B.; Otting, G.; Schlarb-Ridley, B.; Canters, G. W.; Ubbink, M. Hydrophobic interactions in a cyanobacterial plastocyanin-cytochrome *f* complex. *J. Am. Chem. Soc.* **2001**, *123*, 10444−10453.
(48) Banci, L.; Bertini, I.; Cavallaro, G.; Giachetti, A.; Luchinat, C.; Parigi, G. Paramagnetism-based restraints for Xplor-NIH. *J. Biomol. NMR* **2004**, *28*, 249−261.
(49) Palmer, A. G., III; Cavanagh, J.; Wright, P. E.; Rance, M. *J. Magn. Reson.* **1991**, *93*, 151−170.
(50) Schleucher, J.; Schwendinger, M.; Sattler, M.; Schmidt, P.; Schedletzky, O.; Glaser, S. J.; Sørensen, O. W.; Griesinger, C. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR* **1994**, *4*, 301−306.
(51) Bartels, C.; Xia, T. H.; Billeter, M.; Güntert, P.; Wüthrich, K. The program XEASY for computer-supported NMR. *J. Biomol. NMR* **1995**, *5*, 1−10.
(52) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.
(53) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1: a new general purpose quantum mechanical molecular model. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(54) Gasteiger, J.; Marsili, M. Iterative Partial Equalization of Orbital Electronegativity-A Rapid Access to Atomic Charges. *Tetrahedron* **1980**, *36*, 3219−3228.
(55) Schwieters, C. D.; Kuszewski, J.; Tjandra, N.; Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **2003**, *160*, 65−73.

(56) Schwieters, C. D.; Clore, G. M. Internal Coordinates for Molecular Dynamics and Minimization in Structure Determination and Refinement. *J. Magn. Reson.* **2001**, *152*, 288−302.
(57) Kleywegt, G. J.; Jones, T. A. *Methods Enzymol.* **1997**, *277*, 208−230.
(58) MacKerell, A. D., Jr.; Wiòrkiewicz-Kuczera, J.; Karplus, M. An All-Atom Empirical Energy Function for the simulation of Nucleic Acids. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.

JM050574K

4.2 Activity of Anchored Human Matrix metalloproteinase-1 Catalytic Domain on Au (111) Surfaces Monitored by ESI-MS [*J. Mass Spectrom*., 2005, 40, 1565-1571]

JOURNAL OF MASS SPECTROMETRY
J. Mass Spectrom. 2005; **40**: 1565–1571
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jms.929

JMS

Activity of anchored human matrix metalloproteinase-1 catalytic domain on Au (111) surfaces monitored by ESI-MS†

Giuseppe Grasso,[1] Roberta D'Agata,[2] Enrico Rizzarelli,[2,3] Giuseppe Spoto,[2,3]* Luca D'Andrea,[4] Carlo Pedone,[4] Andrea Picardi,[4] Alessandra Romanelli,[5] Marco Fragai[6,7] and Kwon Joo Yeo[6]

[1] Consorzio Interuniversitario di Ricerca in Chimica dei Metalli nei Sistemi Biologici, Via C. Ulpiani 27, 70126, Bari, Italy
[2] Dipartimento di Scienze Chimiche, Università di Catania, Viale Andrea Doria 6, 95125, Catania, Italy
[3] Istituto Biostrutture e Bioimmagini, CNR, Viale A. Doria 6, Catania, Italy
[4] Istituto Biostrutture e Bioimmagini, CNR, Via Mezzocannone 16, Napoli, Italy
[5] Dipartimento delle Scienze Biologiche, Università degli Studi di Napoli ''Federico II'' - Via Mezzocannone 16, Napoli, Italy
[6] Magnetic Resonance Center (CERM), University of Florence, Via L. Sacconi, 6, Sesto Fiorentino, Italy
[7] Department of Agricultural Biotechnology, University of Florence, P.le delle Cascine, 24, Florence, Italy

Received 24 May 2005; Accepted 11 August 2005

Matrix metalloproteinases (MMPs) are a family of Zn-dependent endo-peptidases known for their ability to cleave several components of the extracellular matrix, but which can also cleave many non-matrix proteins. There are many evidences that MMPs are involved in physiological and pathological processes, and a huge effort has been put in the development of possible inhibitors that could reduce the activity of MMPs, as it is clear that the ability to monitor and control such activity plays a pivotal role in the search for potential drugs aimed at finding a cure for several diseases such as pulmonary emphysema, rheumatoid arthritis, fibrotic disorders and cancer.

A powerful method currently available to study enzyme–inhibitor interactions is based on the use of the surface plasmon resonance (SPR) technique. When MMP interactions are studied, a procedure by which inhibitors are normally anchored on sensor chips and SPR technique is used in order to study their interaction with MMPs molecules is usually followed. This is because it is currently believed that MMPs cannot be anchored on the sensor-chip surface without losing their activity. However, this approach gives rise to problems, as the anchoring of low-molecular-weight inhibitors on gold surfaces easily affects their ability to interact with MMPs. For this reason, the anchoring of MMPs is highly desirable.

A new experimental protocol that couples the Fourier transform-SPR (FT-SPR) technique with electrospray ionization-mass spectroscopy (ESI-MS) is described here for the evaluation of the activity of MMP-1 catalytic domain (cdMMP-1) anchored on gold surfaces. The cdMMP-1 surface coverage is calculated by using FT-SPR and the enzyme activity is estimated by ESI-MS. The proposed method is label-free.

KEYWORDS: matrix metalloproteinases; human fibroblast collagenase; activity assay; ESI-MS; SPR

INTRODUCTION

Matrix metalloproteinases[1] (MMPs) are a family of zinc-dependent endo-peptidases known for their ability to cleave several components of the extracellular matrix,[2] but they can also cleave many non-matrix proteins.[3] Human fibroblast collagenase (matrix metalloproteinase-1, MMP-1) was the first vertebrate collagenase both purified to homogeneity as a protein and cloned as a cDNA.[4,5,6,7] It participates in the turnover of collagen fibrils in the extracellular space, but also plays a role in the regulation of cellular behaviour. Furthermore, MMP-1 plays an important role in diverse physiological processes such as tissue morphogenesis and wound repair. Likewise, it seems to be implicated in a variety of human diseases including cancer,[8] rheumatoid arthritis, pulmonary emphysema and fibrotic disorders.[9]

As the role of MMPs in diseases has become better understood, interest in the control of their activity has increased. While *in vivo* the degrading actions of the MMPs are limited by the tissue inhibitors of metalloproteinase (TIMP) family of natural macromolecular inhibitors, it is now clear that MMPs are often overexpressed in diseases such as cancer. A huge effort has been put towards the development of inhibitors able to selectively target MMPs[10,11]

*Correspondence to: Giuseppe Spoto, Dipartimento di Scienze Chimiche, Università di Catania, Viale Andrea Doria 6, 95125, Catania, Italy. E-mail: gspoto@dipchi.unict.it
†Paper presented at the 23rd Informal Meeting on Mass Spectrometry, Fiera di Primiero, Italy, 15–19 May 2005.
Contract/grant sponsor: MIUR; Contract/grant number: (FIRB 2003 RBNE01TTJW$_{004}$ and PRIN 2003 n 2003091372).

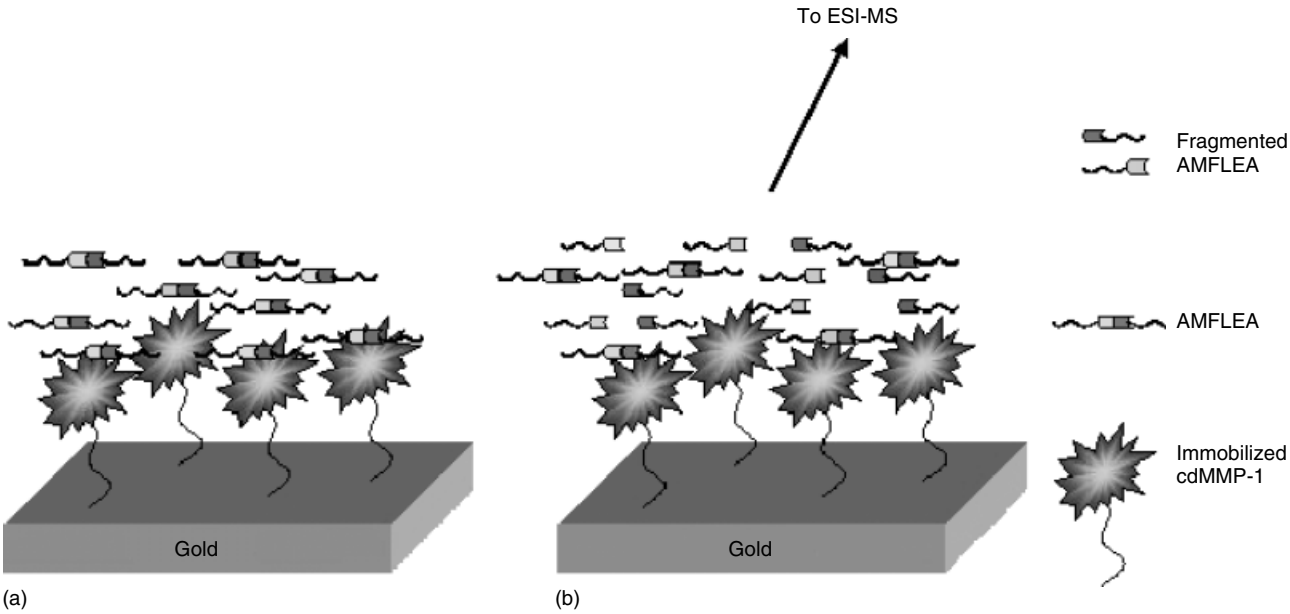


Figure 1. Schematic presentation of the proposed experimental protocol. (a) A gold chip functionalized with dithiobis succinimidylpropionate was used to immobilize the cdMMP-1 molecules. A solution of the cdMMP-1 substrate AMFLEA was then put in contact with the modified chip surface. (b) After appropriate time intervals (as described in the text), portions of the solution in contact with the surface, negligible in volume with respect to the total volume used for the experiment (ratio in the range of 1 : 50), were sampled and analysed by ESI-MS to quantify the amount of fragmented substrate.

for therapeutic intervention in a variety of pathological events. It is clear then that the ability to monitor MMP activity plays a pivotal role in the search for possible drugs.[12]

The current methods to monitor the activity of MMPs are mostly based on spectrophotometric assays[13,14,15] and they have to be carried out in solution. Such methods suffer from inconveniences such as the inner filter correction[16] and the impossibility to recycle the MMPs for different analyses. Moreover, they require a fluorescent tag to be present in the substrate.

To our knowledge, only very few attempts have been made towards the development of high-throughput, low-volume enzyme assays on solid support[17,18,19,20] and none of them has been used to evaluate MMPs' activity. Recently, a solid-phase assay for analysis of MMPs' activity has been proposed but fluorescence tag in the substrate was still necessary to detect the activity.[21] Although it is clear now that coupling different techniques, such as surface plasmon resonance (SPR) and mass spectrometry (MS), is a very powerful way of studying biological molecules,[22,23] only very recently a new strategy for identifying general enzyme inhibitors and monitoring enzyme activity by SPR combined with MS has been demonstrated.[24]

Furthermore, every attempt to study MMPs by SPR has been carried out by anchoring the possible inhibitor and keeping the MMP molecules in the liquid phase,[25] as it is commonly believed that 'immobilization of MMPs is unfavorable because of incorrect orientation of most of the enzyme'.[26] We proved that this last assertion is incorrect as anchoring the MMP-1 molecules on gold surface is not only possible but also very convenient, as some MMPs inhibitors bind to the gold surface with the same chemical groups that should interact with the enzyme.

We developed a new experimental protocol to anchor the catalytic domain of human MMP-1 (hereafter cdMMP-1) on gold surface, preserving its activity. We used the Fourier transform- surface plasmon resonance (FT-SPR) technique to calculate cdMMP-1 gold surface coverage and electrospray ionization-mass spectrometry (ESI-MS) to monitor cdMMP-1 activity (Fig. 1). To our knowledge, this is the first time that cdMMP-1 activity is monitored in a quantitative fashion by ESI-MS.

We believe that this experimental protocol can be widely applied to a vast range of biological systems so that the enzyme is anchored on a gold surface, its interaction with possible inhibitors/substrates can be studied by SPR and its activity can be monitored by ESI-MS. We reasoned that an assay format providing a large contact area between the enzyme and the substrate might give the possibility of monitoring the interaction of some systems that cannot be studied otherwise, as it is in the case of interfacial and low-solubility enzymes.[17]

EXPERIMENTAL

Expression and purification of human cdMMP-1

The cDNA of proMMP-1 (Pro21-Pro269) was cloned into the pET21 vector (Novagen) using *Nde I* and *Xho I* as restriction enzymes. One additional methionine at position 20 was present in the final expression product. The *Escherichia coli* strain BL21 Codon Plus cells transfected with the above vector were grown in 2 × YT media at 37 °C. The protein expression was induced during the exponential growth phase with 0.5 mM of Isopropyl β-D-1-Thiogalactopyranoside (IPTG). Cells were harvested for 4 h after induction.[15]N-labeled protein was expressed in minimal media at 37 °C. After lysis of the cells, the inclusion bodies containing the proMMP-1 were

solubilized in 2 M urea; 20 mM Tris (pH 8.0). The protein was purified on the Hitrap Q column (Pharmacia) with a buffer containing 2 M urea; 20 mM Tris (pH 8.0). The elution was performed using a linear gradient of NaCl up to 0.35 M. The purified protein was then refolded by using a multi-step dialysis against solutions containing 50 mM Tris (pH 7.2), 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$ and 0.3 M NaCl. The refolded protein was exchanged, by dialysis, against a buffer with 10 mM Tris (pH 7.2), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$ and 0.3 M NaCl. The protein was activated by 1 mM APMA (4-aminophenylmercuric acetate) at 4 °C overnight and dialysed with a buffer containing 10 mM Tris (pH 7.2), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl, and 0.2 M acetohydroxamic acid (AHA). The activated protein (Val101-Pro269) was concentrated using an Amicon stirrer and Centriprep concentrator fitted with a YM10 membrane in nitrogen atmosphere at 4 °C. Catalytic domain of MMP-1 was purified using size-exclusion chromatography with the final dialysis buffer and concentrated using a Centriprep concentrator at 4 °C.

Anchoring of cdMMP-1 on gold surface

The anchoring of enzyme molecules on gold surface is usually done in several steps by using a carboxymethyl dextran matrix.[26,27] In our experimental protocol, cdMMP-1 was anchored on the gold surface in just two steps. Firstly, bare gold chips were incubated with dithiobis succinimidylpropionate (Lomant's Reagent) (0.0097 g in 2 ml of dimethyl sulfoxide (DMSO), Sigma-Aldrich) for 48 h. This step was necessary in order to have several succinimidyl ester groups on the surface that could react with the terminal amino groups of the protein that we wanted to be bonded to the surface.[28] The surface so treated was washed with pure DMSO and high-purity water (Milli-Q Element Ultrapure Water) and could react directly with cdMMP-1. We found that the choice of the running buffer (also used for sample dilution) is crucial for the positive result of the whole experiment. The best buffer to be used in order to anchor the cdMMP-1 molecules to the gold surface, preserving their activity, is *N*-(2-hydroxyethyl)piperazine-*N*′-(2-ethanesulfonic acid) sodium salt (0.01 M HEPES, 0.15 M NaCl, pH 7.4; HBS-N, Biacore International AB), indicated as HBS-N buffer hereafter.

cdMMP-1 was anchored on the sensor-chip surface previously modified with the Lomant's reagent by leaving the surface in contact with a 4.5 μM solution of cdMMP-1 in HBS-N buffer for 10 min.

Two different peptides were used as cdMMP-1 substrate. The first one does not have a fluorescent group and it has the sequence AMFLEA (Ac-Ala-Met-Phe-Leu-Glu-Ala-$CONH_2$). It was synthesized on solid phase using Rink Amide MBHA resin (Novabiochem, Switzerland) with standard Fmoc (*N*-(9-fluorenyl)methoxycarbonyl) chemistry. Cleavage from the resin was achieved by treatment with trifluoracetic acid, triisopropyl silane and water (95; 2.5; 2.5) at room temperature for 3 h. Purity and identity of the peptides were assessed by HPLC and MALDI-ToF MS. The second peptide was bought from Biomol International L.P. (PA, USA) and is a known fluorogenic substrate for MMPs having the sequence Ac-Pro-Leu-Gly-[2-mercapto-4-methyl-pentanoyl]-Leu-Gly-OC_2H_5 (P125).[29]

N-Isobutyl-*N*-(4-methoxyphenylsulfonyl)glycyl hydroxamic acid (NNGH) purchased from Biomol International L.P. (PA, USA) was used as cdMMP-1 inhibitor.[30]

Fourier transform- surface plasmon resonance

FT-SPR experiments were carried out by using an FT-SPR 100 (GWC Instruments, USA) apparatus. The light beam from an external port of a Nexus 870 FT-IR spectrometer (Nicolet, USA), equipped with a quartz-halogen source and an XT-KBr beam splitter, was used as the near-IR source of the FT-SPR. Gold substrates (GWC Instruments, USA) were obtained by thermally evaporating a gold layer (450 Å) on to SF-10 glass slides (Schott, USA). Chromium (50 Å) was used as the adhesion layer. Gold substrates were brought into optical contact with the SF-10 equilateral prism present in the FT-SPR by using a refractive index matching fluid (Cargille Laboratories, USA). FT-SPR experiments were carried out by using a 60-μl flow cell (GWC Technologies, USA) and a Masterflex L/S (Cole-Parmer, USA) peristaltic pump.

The FT-SPR sensor response was converted into refractive index changes by using independent sucrose aqueous calibrating solutions ($r^2 = 0.999$). The refractive indices of the sucrose solutions were obtained from the literature.[31]

cdMMP-1 surface coverage

In order to estimate the number of cdMMP-1 molecules anchored on 1 cm^2 of gold surface, we followed the method described in the literature[32] where the adlayer thickness (d) is obtained from the measured SPR response (R), that is the shift in wavenumber of the FT-SPR minimum in reflected light intensity associated with changes in the index of refraction of the medium in contact with the metal surface of the FT-SPR device.

The calculation requires that both the refractive index of the buffer and that of the adsorbate to be known. For this purpose, the refractive index of the buffer was measured with an ABBE-3L refractometer, which provided the value of 1.3349 for the HBS-N buffer, while the refractive index of the adsorbate (the pure cdMMP-1 protein in our case) was taken to be 1.57.[32,33]

About 2 ml of a 4.5 μM solution of cdMMP-1 in HBS-N buffer (this concentration was chosen in order to have the right cdMMP-1 surface coverage for the activity measurements) was eluted in the FT-SPR flow cell at a flow rate of 0.1 ml/min in order to immobilize the enzyme on the sensor-chip surface previously modified with the Lomant's reagent. The flowing HBS-N buffer was then introduced into the flow cell. The FT-SPR shift due to the anchoring of the cdMMP-1 was determined (Fig. 2) and a value for the adlayer thickness $d = 0.84$ nm was obtained. It is then straightforward[34,35] to convert the adlayer thickness into the surface concentration, θ, in molecules per square centimeter:

$$\theta(\text{molecules/cm}^2) = d(\text{cm}) \times N(\text{molecules/cm}^3) \quad (1)$$

where N is the bulk number density of the adsorbate and can be estimated from the bulk density of the adsorbate, ρ, in units of g/cm^3, just by dividing by the molecular weight and multiplying by Avogadro's number. In our case, $\rho = 1.30$ g/cm^3 was obtained from literature[32] and $\theta = 3.6 \times 10^{12}$ (molecules/cm^2) was calculated by using Eqn (1).

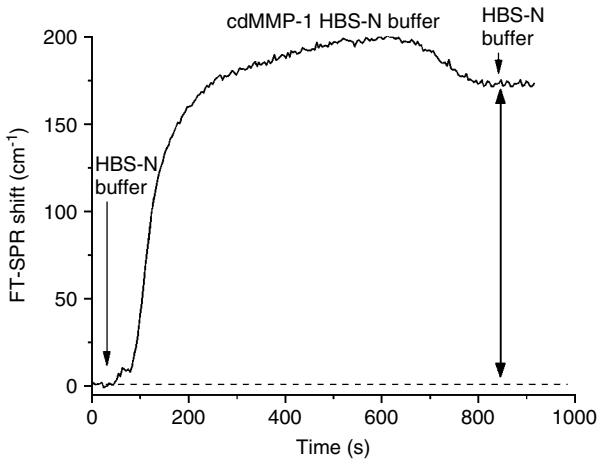


Figure 2. FT-SPR sensorgram for cdMMP-1 immobilization on the sensor-chip surface according to the following steps: 0–80 s: HBS-N buffer solution; 80–700 s: 2 ml of cdMMP-1 in HBS-N buffer (4.5 μM); 700–900 s: HBS-N buffer solution. Flow-rate 0.1 ml/min. The double head arrow indicates the FT-SPR response that is proportional to the amount of anchored cdMMP-1.

Electrospray ionization-mass spectrometry (ESI-MS)

All the ESI-MS measurements were carried out by using a Finnigan LCQ DECA XP PLUS ion trap spectrometer operating in the positive ion mode and equipped with an orthogonal ESI source (Thermo Electron Corporation, USA). Sample solutions were injected into the ion source at a flow-rate of 10 μl/min, using nitrogen as the drying gas. The mass spectrometer operated with a capillary voltage of 46 V and a capillary temperature of 200 °C, while the spray voltage was 4.3 kV.

Xcalibur software was used for the elaboration of mass spectra and for the quantitative evaluation of the MS data.

According to several other quantitative studies,[36,37] the unavailability of an internal standard for the analysis can be overcome by the use of a calibration curve. To prepare stock standard solutions, 0.0006 g of AMFLEA was accurately weighed and dissolved in 1 ml of HBS-N buffer to generate a solution of 914 μM that was further diluted in order to generate calibration standards, covering the concentration range 10–600 μM. Ten calibration standards were used to define the calibration curve (10, 100, 130, 151, 230, 270, 350, 402, 450 and 503 μM) (Fig. 3). Four quality control (QC) standards (100, 151, 270 and 402 μM) were also used for intra- and inter-day validation.

It was noted that subtle changes in the amount of NaCl dissolved in the buffer cause appreciable changes in the amount of AMFLEA detected by ESI-MS.[38] In order to obtain high-sensitivity spectra, peptide solutions were injected as 25 μl aliquots into the ESI-MS apparatus and data were recorded in SIM mode (centre m/z 766 and width 50 Da for AMFLEA). The peptide was represented in the ESI-MS positive ion spectra by two peaks at m/z 744.5 and 766.4 (mono- and bi-sodiate respectively). Quantitation of the peptide concentration was carried out by monitoring the most intense peak at m/z 766.4, and all calibration curves showed good

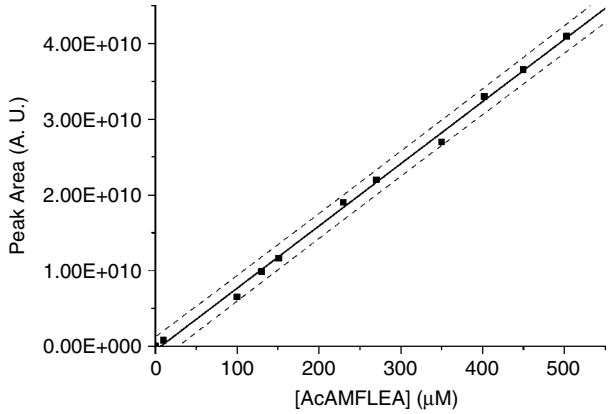


Figure 3. ESI-MS calibration curve for AMFLEA. The ion at m/z 766.4 was selected for SIM mode (centre 766, width 50). The solid line results from the linear regression of the MS data ($r^2 = 0.998$), while the dashed lines represent the upper and lower 95% prediction limits.

linearity ($r^2 > 0.99$) over the concentration range. By setting the base peak at m/z 766 it was possible to obtain an ion chromatogram that presented a shape approaching a rectangle, the integral of which was proportional to the concentration of the peptide solution injected.

The results were reproducible only if the buffer solution had always the same amount of NaCl dissolved. For this reason, we always used the same buffer and gave particular attention in order to avoid solvent evaporation.

RESULTS AND DISCUSSION

A critical aspect of the cdMMP-1 activity assay described here was represented by the proper preparation and maintenance of the modified gold chip surface. In particular, it was found that a too dense packing of the anchored cdMMP-1 affected the enzyme activity while lower surface concentration of cdMMP-1 was not suitable for the reliable ESI-MS detection of the substrate used in the enzyme kinetics assay. The negative influence of higher surface packing of uninhibited cdMMP-1 has been attributed to the autolytic cleavage of the enzyme molecules.[39] Another critical point arises from the necessity to keep the cdMMP-1 anchored on the surface constantly wet in order to preserve the enzyme activity (data not shown).

In our approach, the quantitative responses of the FT-SPR shift with solutions of known concentration of cdMMP-1 were used to estimate the enzyme coverage of the gold surface (Experimental Section). We found that a surface coverage of 3.6×10^{12} molecules/cm^2 is adequate for an ESI-MS kinetics assay of cdMMP-1 anchored on gold chips.

The positive ion ESI-MS spectrum of the AMFLEA substrate peptide (MW = 722 Da) showed a peak at m/z 766.4 attributed to the bi-sodiated $[AMFLEA\text{-}H + Na_2]^+$ ion species and a peak at m/z 744.5 attributed to the mono-sodiated $[AMFLEA + Na]^+$ peptide ions. (Fig. 4(a)) When AMFLEA solutions were left in contact with the cdMMP-1 anchored onto the gold chip surfaces, new species were generated. The cdMMP-1 cleavage of the substrate generates two peptide fragments[40] detectable by ESI-MS at m/z

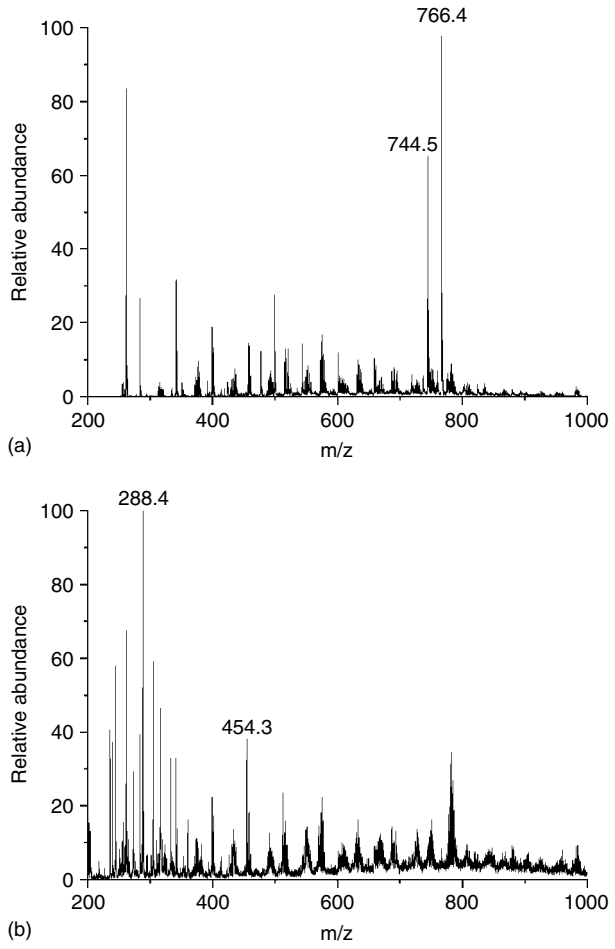


Figure 4. (a) Representative positive ion ESI-MS spectrum of the AMFLEA peptide in HBS-N buffer (flow rate 10 μl/min). The peaks of the peptide at *m/z* 766.4 and 744.5 are visible together with many other peaks of the HBS-N buffer. (b) Representative positive ion ESI-MS spectrum of the AMFLEA peptide in HBS-N buffer after incubation for 1 h with cdMMP-1. The peaks at *m/z* 766.4 and 744.5 disappear and two new peaks at *m/z* 454.3 and 288.4 are now present. Explanation is in the text.

454.3 $[AcAMF - Na_2]^+$ and at m/z 288.4 $[LE\text{-}NHCHCH_3]^+$ (Fig. 4(b)). Unfortunately, many peaks coming from the HBS-N buffer are found to be around the 200–500-Da range, so that working quantitatively in this range suffers from interference from the buffer. For this reason, ESI-MS data for the quantitative evaluation of the progressing of the cdMMP-1/substrate reaction were based on the decrease in the m/z 766.4 peak intensity (SIM mode, centre m/z 766, width 50 Da).

One-hundred-microlitre aliquots having three representative concentrations of AMFLEA (100, 250 and 400 μM) in HBS-N buffer were put in contact with the cdMMP-1 modified sensor-chip surface, and 2-μl aliquots were sampled every minute. Every aliquot was then added with 450 μl of methanol, and 25 μl of the resultant solution was injected into the ESI-MS apparatus. This large dilution of the original aliquot was found to be necessary in order to avoid experimental problems such as capillary obstruction and high-spray current values.

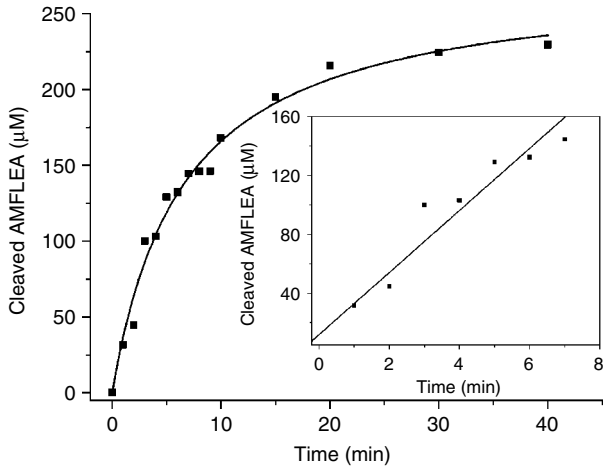


Figure 5. The decrease of the initial concentration of AMFLEA (250 μM was the initial concentration in this graph) is used to calculate the increased amount of AMFLEA cleaved by cdMMP-1. The cdMMP-1 activity was thus monitored, and it is possible to see that for this concentration and for time <10 min the curve can be approximated to a straight line (as shown in the insert).

The calibration curve generated from the standard solutions (Experimental Section) allowed us to convert the m/z 766.4 peak intensity values into peptide concentration ([AMFLEA]) of the different aliquot samples, so that we were able to obtain progress curves of the reaction, such as the one shown in Fig. 5. The progress curves were obtained by plotting the amount of cleaved AMFLEA *versus* time and were used to determine the optimum enzyme concentration and reaction time for the kinetic study. From Fig. 5, it is possible to verify that for a reaction time <10 min the curve obtained can be approximated to a straight line. An optimal reaction time of 5 min was determined. Since the chosen optimal reaction time resided in the linear region of the progression curve, the initial velocity (V_i) was simply calculated by dividing the difference between the ESI-MS determined AMFLEA amounts at $t = 0$ min and $t = 5$ min by the reaction time. The values of V_i for different initial concentrations of AMFLEA in contact with the cdMMP-1 chip were also obtained in the same way.[41] It was previously demonstrated that MMPs follow a Michaelis–Menten kinetics,[42] so the K_M and V_{max} values can be estimated from the double-reciprocal plot of $1/V_i$ *vs* 1/[AMFLEA]. A representative Lineweaver–Burk plot for the cdMMP-1 on chip/AMFLEA reaction is shown in Fig. 6.

Three independent measurements of cdMMP-1 activity were carried out and the average values of $K_M = 3.5\ (\pm 2) \times 10^3$ (μM) and $V_{max} = 1.8 \pm (0.2) \times 10^2$ (μM/min) were obtained. The high error affecting the K_M value is mainly attributed to uncontrolled dilution of the initial substrate concentration in contact to the enzyme chip surface caused by the necessity to avoid the drying of the surface. The drying of the cdMMP-1 chip surface was observed to significantly affect enzyme activity. Nevertheless, it can be seen that the K_M and V_{max} values obtained are relatively large in comparison with the values reported in the literature for MMP-1/substrate interaction in solution.[43]

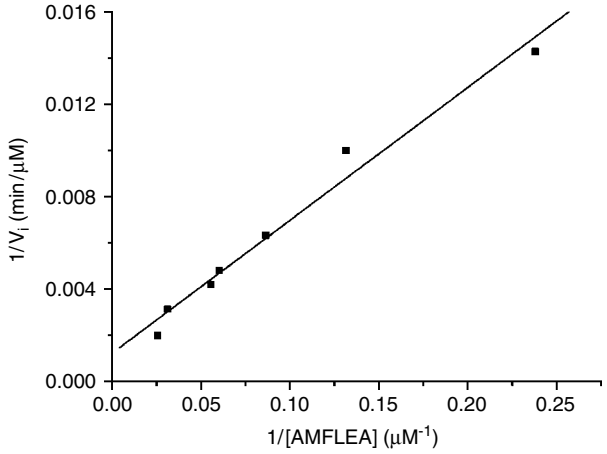


Figure 6. Lineweaver–Burk plot for cdMMP-1/AMFLEA at room temperature ($r^2 = 0.977$).

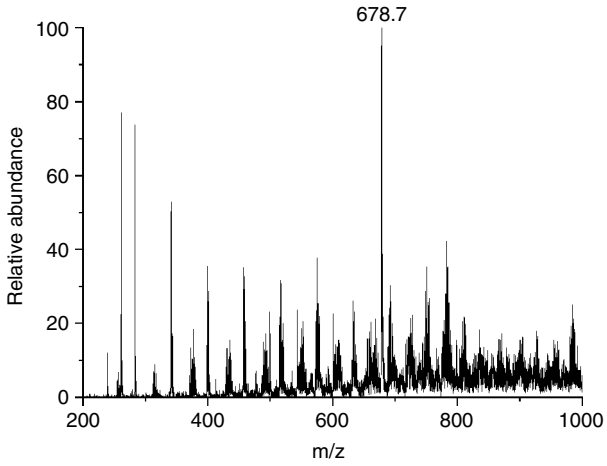


Figure 7. ESI-MS spectrum of P125 in HBS-N buffer, full scan range 200–1000 Da. The peak at *m/z* 678.7, attributed to the mono-sodiated ion species, is visible together with many other peaks generated from species present in the HBS-N buffer.

This is due to the higher number of enzyme molecules per substrate molecules required for the proposed activity assay compared to the usual enzyme/substrate ratio used in spectrophotometric methods. Nevertheless, the kinetic results can be compared with those reported in the literature by looking at the k_{cat} value. This last parameter takes into account the enzyme concentration ($k_{cat} = V_{max}/E_0$, where E_0 is the enzyme concentration) and it is possible to note that the k_{cat} ($= 4.6 \pm 0.6\ sec^{-1}$) value obtained from our experiments is similar to those reported in the literature for similar systems.[43]

By following the described procedure, kinetic assay was also carried out by using the commercially available P125 substrate (MW = 656 Da). In this case, kinetic analysis was monitored by following the intensity of the mono-sodiated peak at *m/z* 678.7, (Fig. 7) and statistically equivalent kinetic parameters were obtained.

The same experimental procedure was carried out after inhibition of cdMMP-1 by NNGH inhibitor. A lack of activity by cdMMP-1 with both substrates (AMFLEA and P125) was observed (Fig. 8). The same result was obtained in two

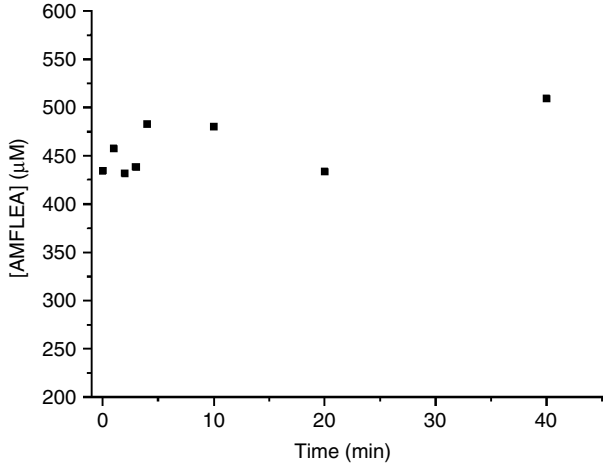


Figure 8. There is no progression in the reaction between NNGH-inhibited cdMMP-1 and the AMFLEA substrate. No appreciable variation with time is detected for the ESI-MS-determined AMFLEA concentration of the substrate solution in contact with NNGH-inhibited cdMMP-1. The graph shown refers to an experiment carried out with a [AMFLEA] = 450 μM.

different cases, where cdMMP-1 incubation with NNGH was carried out prior to or after the enzyme immobilization.

CONCLUSIONS

A new experimental protocol was developed that enabled us to anchor cdMMP-1 enzyme on gold surface. The immobilization did not affect cdMMP-1 activity, which could be monitored by ESI-MS. The feasibility of this approach was demonstrated for two different cdMMP-1 substrates, and inhibition with NNGH of the anchored cdMMP-1 was also proved.

These results demonstrate the possibility of monitoring cdMMP-1 inhibition on solid-state support. By coupling the FT-SPR and ESI-MS techniques, a comprehensive study of the kinetics of inhibition and of the possibility to determine the activity of MMPs is shown to be feasible.

We believe that the new solid-state activity assay opens the doors to a wide range of applications where the activity of an enzyme has to be monitored in the presence of possible inhibitors. In our new experimental protocol, substrates do not need to have fluorescent groups and one requires only a small amount of enzyme to carry out several different screenings.

Other MMPs together with other enzymes are currently being studied in our laboratory in order to expand the applicability of this new methodology. Efforts are also on to improve accuracy and precision of the assay.

Acknowledgement

We thank MIUR (FIRB 2003 RBNE01TTJW$_{004}$ and PRIN 2003 n 2003091372) for partial financial support.

REFERENCES

1. Massova I, Kotra LP, Fridman R, Mobashery S. Matrix metalloproteinases: structures, evolution, and diversification. *FASEB J.* 1998; **12**: 1075.

2. Christopher M. Molecular determinants of metalloproteinase substrate specificity: matrix metalloproteinase substrate binding domains, modules, and exosites. *Mol. Biotechnol.* 2002; **22**: 51.
3. Nagase H, Woessner JF. Matrix metalloproteinases. *J. Biol. Chem.* 1999; **274**: 21491.
4. Pardo A, Selman M. MMP-1: the elder of the family. *Int. J. Biochem. Cell Biol.* 2005; **37**: 283.
5. Bauer EA, Eisen AZ, Jeffrey JJ. Immunologic relationship of a purified human skin collagenase to other human and animal collagenases. *Biochim. Biophys. Acta* 1970; **206**: 152.
6. Goldberg GI, Wilhelm SM, Kronberger A, Bauer EA, Grant GA, Eisen AZ. Human fibroblast collagenase. Complete primary structure and homology to an oncogene transformation-induced rat protein. *J. Biol. Chem.* 1986; **261**: 6600.
7. Lovejoy B, Hassell AM, Luther MA, Weigl D, Jordan SR. Crystal structures of recombinant 19-kDa human fibroblast collagenase complexed to itself. *Biochemistry* 1994; **33**: 8207.
8. Coussens LM, Werb Z. Matrix metalloproteinases and the development of cancer. *Chem. Biol.* 1996; **3**: 895.
9. Liu Z, Zhou XY, Shapiro SD, Shipley JM, Twining SS, Diaz LA. The serpin alpha1-proteinase inhibitor is a critical substrate for gelatinase B/MMP-9 in vivo. *Cell* 2000; **102**: 647.
10. Farkas E, Katz Y, Bhusare S, Reich R, Röschenthaler G, Königsmann M, Breuer E. Carbamoylphosphonate based matrix metalloproteinase (MMP) inhibitor metal complexes – solution studies and stability constants. Towards a Zinc Selective Binding Group. *J. Biol. Inorg. Chem.* 2004; **9**: 307.
11. Koivunen E, Arap W, Valtanen H, Rainisalo A, Medina OP, Heikkilä P, Kantor C, Gahmberg1 CG, Salo T, Konttinen YT, Sorsa T, Ruoslahti E, Pasqualini R. Tumor targeting with a selective gelatinase inhibitor. *Nat. Biotechnol.* 1999; **17**: 768.
12. Dive V, Andarawewa KL, Boulay A, Matziari M, Beau F, Guerin E, Rousseau B, Yiotakis A, Rio M. Dosing and scheduling influence the antitumor efficacy of a phosphinic peptide inhibitor of matrix metalloproteinases. *Int. J. Cancer* 2005; **113**: 775.
13. Xia T, Akers K, Eisen AZ, Seltzer JL. Comparison of cleavage site specificity of gelatinases A and B using collagenous peptides. *Biochim. Biophys. Acta* 1996; **1293**: 259.
14. Weingarten H, Feder J. Spectrophotometric assay for vertebrate collagenase. *Anal. Biochem.* 1985; **147**: 437.
15. Beekman B, Drijfhout JW, Bloemhoff W, Ronday HK, Tak PP, te Koppele JM. Convenient fluorometric assay for matrix metalloproteinase activity and its application in biological media. *FEBS Lett.* 1996; **390**: 221.
16. Liu Y, Kati W, Chen C, Tripathy R, Molla A, Kohlbreuner W. Use of a fluorescence plate reader for measuring kinetic parameters with inner filter effect correction. *Anal. Biochem.* 1999; **267**: 331.
17. Babiak P, Reymond J. A high-throughput, low-volume enzyme assay on solid support. *Anal. Chem.* 2005; **77**: 373.
18. Min D, Mrksich M. Profiling kinase activities by using a peptide chip and mass spectrometry. *Angew. Chem. Int. Ed. Engl.* 2004; **43**: 5973.
19. Houseman BT, Huk JH, Kron SJ, Mrksich M. Peptide chips for the quantitative evaluation of protein kinase activity. *Nat. Biotechnol.* 2002; **20**: 270.
20. Huber A, Demartis S, Neri D. The use of biosensor technology for the engineering of antibodies and enzymes. *J. Mol. Recognit.* 1999; **12**: 198.
21. Lauer-Fields JL, Nagase H, Fields GB. Development of a solid-phase assay for analysis of matrix metalloproteinase activity. *J. Biomol. Tech.* 2004; **15**: 305.
22. Nedelkov D, Nelson RW. Analysis of native proteins from biological fluids by biomolecular interaction analysis mass spectrometry (BIA/MS): exploring the limit of detection, identification of non-specific binding and detection of multi-protein complexes. *Biosens. Bioelectron.* 2001; **16**: 1071.
23. Nedelkov D, Nelson RW. Pratical considerations in BIA/MS: optimizing the biosensor-mass spectrometry interface. *J. Mol. Recognit.* 2000; **13**: 140.
24. Borch J, Roepstorff P. Screening for enzyme inhibitors by surface plasmon resonance combined with mass spectrometry. *Anal. Chem.* 2004; **76**: 5243.
25. Olson MW, Gervasi DC, Mobashery S, Fridman R. Kinetic analysis of the binding of human matrix metalloproteinase-2 and -9 to tissue inhibitor of metalloproteinase (TIMP)-1 and TIMP-2. *J. Biol. Chem.* 1997; **272**: 29975.
26. Pieper-Fürst U, Kleuser U, Stöcklein WFM, Warsinke A, Scheller FW. Detection of subpicomolar concentrations of human matrix metalloproteinase-2 by an optical biosensor. *Anal. Biochem.* 2004; **332**: 160.
27. Johnsson B, Löfas S, Lindquist G. Immobilization of proteins to a carboxy methyldextran-modified gold surface for biospecific interaction analysis insurface plasmon resonance sensors. *Anal. Biochem.* 1991; **198**: 268.
28. Dordi B, Schönherr H, Vancso GJ. Reactivity in the confinement of self-assembled monolayers: chain length effects on the hydrolysis of *N*-hydroxysuccinimide ester disulfides on gold. *Langmuir* 2003; **19**: 5780.
29. Weingarten H, Martin R, Feder J. Synthetic substrates of vertebrate collagenase. *Biochemistry* 1985; **24**: 6730.
30. MacPherson LJ, Bayburt EK, Capparelli MP, Carroll BJ, Goldstein R, Justice MR, Zhu L, Hu S, Melton RA, Fryer L, Goldberg RL, Doughty JR, Spirito S, Blancuzzi V, Wilson D, O'Byrne EM, Ganu V, Parker DT. Discovery of CGS 27023A, a non-peptidic, potent, and orally active stromelysin inhibitor that blocks cartilage degradation in rabbits. *J. Med. Chem.* 1997; **40**: 2525.
31. Lide DR. *Handbook of Chemistry and Physics*, 71st ed., CRC Press: Boston, 1990.
32. Jung LS, Campbell CT, Chinowsky TM, Mar MN, Yee SS. Quantitative interpretation of the response of surface plasmon resonance sensors to adsorbed films. *Langmuir* 1998; **14**: 5636.
33. Armstrong SH, Budkak MJE, Morrison C, Hasson M. Preparation and properties of serum and plasma proteins. XII. The refractive properties of the proteins of human plasma and certain purified fraction. *J. Am. Chem. Soc.* 1947; **69**: 1747.
34. Darnell JE, Lodish H, Baltimore D. *Molecular Cell Biology*. Scientific American Books: New York, 1990.
35. Leslie TE, Lilley TH. Aqueous solutions containing amino acids and peptides. Part 20. Volumetric behavior of some terminally substituted amino acids and peptides at 298.15 K. *Biopolymers* 1985; **24**: 695.
36. Wang X, Sakuma T, Asafu-Adjaye E, Shiu GK. Determination of ginsenosides in plant extracts from Panax ginseng and Panax quinquefolius L. By LC/MS/MS. *Anal. Chem.* 1999; **71**: 1579.
37. Mauri P, Migliazza B, Pietta P. Liquid chromatography/electrospray mass spectrometry of bioactive terpenoids in Ginkgo biloba L. *J. Mass Spectrom.* 1999; **34**: 1361.
38. Iavarone AT, Udekwu OA, Williams ER. Buffer loading for counteracting metal salt-induced signal suppression in electrospray ionization. *Anal. Chem.* 2004; **76**: 3944.
39. Clark IM, Mitchell RE, Powell LK, Bigg HF, Cawston TE, O'Hare MC. Recombinant porcine collagenase: purification and autolysis. *Arch. Biochem. Biophys.* 1995; **316**: 123.
40. Neumann U, Kubota H, Frei K, Ganu V, Leppert D. Characterization of Mca-Lys-Pro-Leu-Gly-Leu-Dppa-Ala-Arg-NH_2, a fluorogenic substrate with increased specificity constants for collagenases and tumor necrosis factor converting enzyme. *Anal. Biochem.* 2004; **328**: 166.
41. Pi N, Armstrong JI, Bertozzi CR, Leary JA. Kinetic analysis of NodST sulfotransferase using an electrospray ionization mass spectrometry assay. *Biochemistry* 2002; **41**: 13283.
42. Nagase H, Fields CG, Fields GB. Design and characterization of a fluorogenic substrate selectively hydrolyzed by stromelysin 1 (matrix metalloproteinase-3). *J. Biol. Chem.* 1994; **269**: 20952.
43. Neumann U, Kubota H, Frei K, Ganu V, Leppert D. Characterization of Mca-Lys-Pro-Leu-Gly-Leu-Dpa-Ala-Arg-NH_2, a fluorogenic substrate with increased specificity constants for collagenases and tumor necrosis factor converting enzyme. *Anal. Biochem.* 2004; **328**: 166.

4.3 In Situ AP/MALDI-MS characterization of anchored MMPs [*J. Mass Spectrom*., 2006, 41, 1561-1569]

JOURNAL OF MASS SPECTROMETRY
J. Mass Spectrom. 2006; **41**: 1561–1569
Published online 9 November 2006 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/jms.1126

JMS

In Situ AP/MALDI-MS characterization of anchored matrix metalloproteinases†

Giuseppe Grasso,[1] Marco Fragai,[2,3] Enrico Rizzarelli,[4,5] Giuseppe Spoto[4,5]* and Kwon Joo Yeo[2]

[1] Consorzio Interuniversitario di Ricerca in Chimica dei Metalli nei Sistemi Biologici, Via C. Ulpiani 27, Bari, Italy
[2] Magnetic Resonance Center (CERM), University of Florence, Via L. Sacconi, 6, Sesto Fiorentino, Italy
[3] Department of Agricultural Biotechnology, University of Florence, P.le delle Cascine, 24, Florence, Italy
[4] Dipartimento di Scienze Chimiche, Università di Catania, Viale Andrea Doria 6, 95125, Catania, Italy
[5] Istituto Biostrutture e Bioimmagini, CNR, Viale A. Doria 6, Catania, Italy

Received 16 May 2006; Accepted 22 September 2006

Several different procedures are available for the immobilization of proteins on solid supports, as many advantages derive from this approach, such as the possibility to develop new protein solid-state assays. Enzymes that are anchored on gold surfaces can interact with several different molecules in a tag-free environment, opening the way to surface plasmon resonance (SPR) investigations. Nevertheless, it is often important to know the identity of the affinity-retained analyte, and mass spectrometric analysis, via its unique molecular mass identification, represents a very valuable complementary method.

There are many pieces of evidence to suggest that matrix metalloproteinases (MMPs) are involved in normal and pathological processes, including embryogenesis, wound healing, inflammation, arthritis and cancer, but presumably also exhibiting other functions. The search for new inhibitors of MMPs has prompted research towards the development of new solid-state assays for the rapid evaluation of MMP activity. We have already reported the possibility of measuring the activity of MMP-1 anchored on solid support by coupling SPR with ESI-MS analysis. In this work, we show the *in situ* atmospheric pressure (AP) MALDI-MS characterization of MMPs anchored on a gold chip with known surface coverage. The study extends the MS analysis to different proteins, and sequence coverage is reported for different digestion and MS procedures. Copyright © 2006 John Wiley & Sons, Ltd.

KEYWORDS: matrix metalloproteinases; AP/MALDI-MS; SPR; solid-state assay; MMP immobilization

INTRODUCTION

Solid-state assays represent a very powerful approach for protein studies,[1,2] and in the past few years some attempts to develop high-throughput and low-volume enzyme assays on solid supports have been carried out.[3–5] This approach offers some important advantages that are not available in the standard solution assays, such as the possibility to recycle the studied enzymes, a tag-free working environment that allows simple and rapid on-spot protein profiling,[6] high throughput,[7] improved discovery of protein binding partners through protein-affinity interactions,[8] and, most importantly, a multiplexed approach for the investigation of proteins, allowing the coupling of different techniques such as, for instance, surface plasmon resonance (SPR) and mass spectrometry (MS).[9–11] The latter multiplexed approach is particularly advantageous in the case of protein arrays[12] and they are poised to become a central proteomics technology important both in basic research and commercially for biotechnology enterprises. In recent years many studies have been addressed toward the development of different experimental protocols aimed at improving the efficiency and applicability of such an approach.[13]

Laser desorption ionization MS-based methods play a central role in the above-mentioned scenario. They do not require analytes to be labeled and therefore are applicable to the study of a broad range of biological molecules. High sensitivity is another typical advantage offered by MS-based methods, while the use of lasers guarantees the spatial resolution of the analysis and the suitability for solid-state assays on the array format.

Surface-enhanced laser desorption ionization (SELDI)[14] is one of the above-mentioned MS-based methods that has seen rapid development[15–17] recently for a broad range of applications.[18–20] The reason for such success lies mainly in the new research opportunities and advantages that SELDI investigations offer for on-chip protein characterization, including on-spot deglycosylation or dephosphorylation, with subsequent identification through peptide mapping or MS/MS sequencing.[21]

Desorption/ionization on silicon mass spectrometry (DIOS-MS)[22] and self-assembled monolayers for MALDI

*Correspondence to: Giuseppe Spoto, Dipartimento di Scienze Chimiche, Università di Catania, Viale Andrea Doria 6, 95125, Catania, Italy. E-mail: gspoto@dipchi.unict.it
†Paper presented at the 24th Informal Meeting on Mass Spectrometry, Ustroń, Poland, 14–18 May 2006.

(SAMDI)[23] offer further MS-based tools for solid-state assays of proteins.

Unfortunately, one of the limiting factors of most of the above-mentioned approaches is the difficulty in obtaining good quality MS data, mostly due to the mixing of the sample of interest with salt and buffer impurities, the presence of which is often necessary when dealing with biological molecules.[24,25] As a direct consequence of this problem, several papers focusing on all the practical considerations in biomolecular interaction analysis mass spectrometry (BIA/MS) have been recently published.[26] Purification of contaminated peptides and proteins,[27,28] choice of matrix molecules and MALDI sample platform used,[29,30] appropriate MALDI spotting procedure,[31] addition of special molecules[32] and interpretation of salt-affected MS spectra[33,34] are all examples that show the difficulties normally encountered in this field.

In a previous paper,[35] we have already mentioned the important role that matrix metalloproteinases (MMPs)[36] have in normal and pathological processes, including embryogenesis, wound healing, inflammation, arthritis and cancer.[37] It is clear that while *in vivo* the degrading actions of MMPs are limited by the tissue inhibitors of metalloproteinase (TIMP) family of natural macromolecular inhibitors, in diseases such as cancer MMPs are often overexpressed.[38] For this reason, only very recently, much effort has been put in the development of inhibitors capable of selectively targeting MMPs for therapeutic intervention in a variety of pathological events.[39] In this perspective, the development of new solid-state assays for a rapid evaluation of MMP activity and characterization is of great interest.

We have already showed the possibility to study the activity of the catalytic domain of MMP-1 (cdMMP-1) by using a solid-state assay based on both Fourier transform (FT)-SPR and electrospray ionization (ESI)-MS techniques.[35] cdMMP-1 was anchored on a pre-functionalized gold surface, and it was demonstrated that the immobilization procedure did not affect the enzyme activity. Although in principle such an approach could be applied to other enzymes, there was no easy way of unambiguously identifying the protein immobilized on the gold surface.

The need for reliable procedures for the characterization of immobilized proteins is strictly related to the development of protein arrays. Protein array fabrication requires different steps to be used. Roughly, (1) proteins are immobilized on a solid surface after reaction with suitable chemical compounds previously anchored to the solid surface, (2) unreacted chemical compounds present on the solid surface are deactivated and (3) regions of the solid surface not involved in the protein immobilization are modified with specific compounds capable of improving the surface resistance to nonspecific interactions. Accurate washing with appropriate solvents are imperative after each step. It is evident that the removal of proteins from each array region and their re-binding in the surrounding areas is possible, thus resulting in a mixing of proteins. The development of protein arrays requires the optimization of the above-mentioned procedures and it is not possible without methods for identifying the proteins immobilized on the gold surface. Moreover, such methods must be able to identify the presence of unwanted mixing caused by the re-binding of proteins coming from close regions of the array. In order to contribute to the development of suitable methods to be used in the characterization of protein arrays, we present a method to characterize anchored MMPs here. The method is based on the experimental procedure of *in situ* digestion and atmospheric pressure (AP) MALDI-MS[40] characterization of the covalently immobilized enzymes. In order to demonstrate the suitability of the method for immobilized MMPs characterization, different cdMMPs (cdMMP-1, cdMMP-8, cdMMP-12) and their mixtures, as well as control samples such as bovine serum albumin (BSA), are used as examples, and the strict control over the experimental procedures used for the immobilization and characterization of the enzyme is highlighted. In order to avoid MMP degradation and deactivation, such procedures require the use of buffers with an appropriate salt content.[35] The proposed approach is able to overcome some experimental difficulties related to the presence of salts and buffers, and the results represent a further example of the high impact that such contaminants can have on the appearance of AP/MALDI-MS spectra.[41]

EXPERIMENTAL

Protein cloning and purification

The synthesis and purification of cdMMP-1 has already been described elsewhere.[35]

The cDNA encoding the cdMMP-12 (Gly106-Gly263) and cdMMP-8 (100M-G262) was amplified by a polymerase chain reaction (PCR) from the IMAGE consortium clone (ID 196 612) and human cDNA clone (ORIGENE), respectively, using two synthetic oligonucleotides as primers. The cDNA obtained was cloned into the pET21a vector (Novagen) between the restriction sites NdeI and BamHI. For cdMMP-12, the single amino acid substitution, to obtain the F171D mutant, was created using the QuickChange Site-directed mutagenesis kit from Stratagene. For both cdMMP-12 and cdMMP-8, the construct was transformed into the BL21Gold (DE3) strain for expression of the recombinant protein. The cells were grown in 2 × YT media at 37 °C. The protein expression was induced during the exponential growth phase with 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested for 4 h after induction. After lysis of the cells, the inclusion bodies containing the protein were solubilized in 8 M urea and 20 mM sodium acetate (pH 5.0 for cdMMP-12, pH 8.0 for cdMMP-8). The protein was purified on the Hitrap (SP for cdMMP-12, Q for cdMMP-8) column (Pharmacia) with a buffer containing 6 M urea, 20 mM sodium acetate (pH 5.0) (cdMMP-12) and 20 mM Tris (pH 8.0) (cdMMP-8). The elution was performed using a linear gradient of NaCl up to 0.35 M. The purified protein was re-folded by using a multistep dialysis against a solution containing 50 mM Tris-HCl (pH = 7.2), 10 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl and 500 mM acetohydroxamic acid (AHA), and decreasing the concentration of urea (from 4 M down to 2 M). The re-folded protein was concentrated up to 50 μM using an Amicon stirrer and Centriprep concentrator, fitted with a YM10 membrane, at 4 °C.

DOI: 10.1002/jms

cdMMP-1 and cdMMP-12 were ^{15}N enriched (95%), and this was taken into account for the MS analysis.

Surface plasmon resonance imaging (SPRI) apparatus

SPRI experiments were carried out at room temperature by using an SPR imager apparatus (GWC Technologies, USA) equipped with a white light source and an SF-10 prism. A narrow bandpass filter (800 nm) was placed before the CCD camera used as a detector for SPR images.

SPR images were analyzed by using the V + + software (version 4.0, Digital Optics Limited, New Zealand). SPRI provides data as pixel intensity units on a 0–255 scale. As specified by the manufacturer, data are converted into percentage of reflectivity (%R), or Δ%R in the case of difference images, by using the formula:

$$\%R = 100 \times (0.85 I_p / I_S)$$

where I_p and I_s refer to the reflected light intensity detected using p- and s-polarized light, respectively.

PDMS microfluidic channels fabrication and protein array

Microfluidic channels were fabricated in poly(dimethylsiloxane) (PDMS) polymer as described elsewhere.[42] Briefly, PDMS microchannels were created by replication from a master in polyvinyl chloride (PVC), with a pattern of three (80 μm depth, 1.4 cm length, 400 μm width) parallel channels, featuring circular reservoirs (diameter = 0.4 mm) at both ends of each channel. PEEK tubes (0.5 mm, UpChurch Scientific) were inserted in such reservoirs in order to connect the PDMS microfluidic cell to a Masterflex L/S (Cole-Parmer, USA) peristaltic pump,[43] operating at 100 μl/min. Replicas were formed from a 1 : 10 mixture of PDMS curing agent and prepolymer (Sylgard 184, Dow Corning, Midland, MI). The mixture was degassed under vacuum and then poured onto the master in order to create a layer with a thickness of about 3–4 mm. The PDMS replica was then cured for at least 2 h at 60 °C before it was removed from the master.

An array of cdMMPs was obtained by flowing three different cdMMP solutions into the three PDMS microchannels in contact with a prefunctionalized gold substrate.[35] In Fig. 1(a) an SPR image of the protein array so formed is shown, while in Fig. 1(b) the change of percent reflectivity over time for the immobilization of cdMMP-1 is also reported as a representative example, where the difference in the signal intensity between the first and the last parts of the graph is proportional to the amount of immobilized enzyme (indicated as a double arrow in Fig. 1(b)).

Surface coverage: FT-SPR measurements

The FT-SPR apparatus and the procedure used to anchor BSA, cdMMP-1, cdMMP-8 and cdMMP-12, as well as the equations used to calculate the protein surface coverage, were the same as described previously.[35] Surface coverage is governed by two main factors: the initial protein concentration and the contact time of the protein solution with the surface. The latter experimental parameters were optimized in order to obtain the best quality MS spectra, so

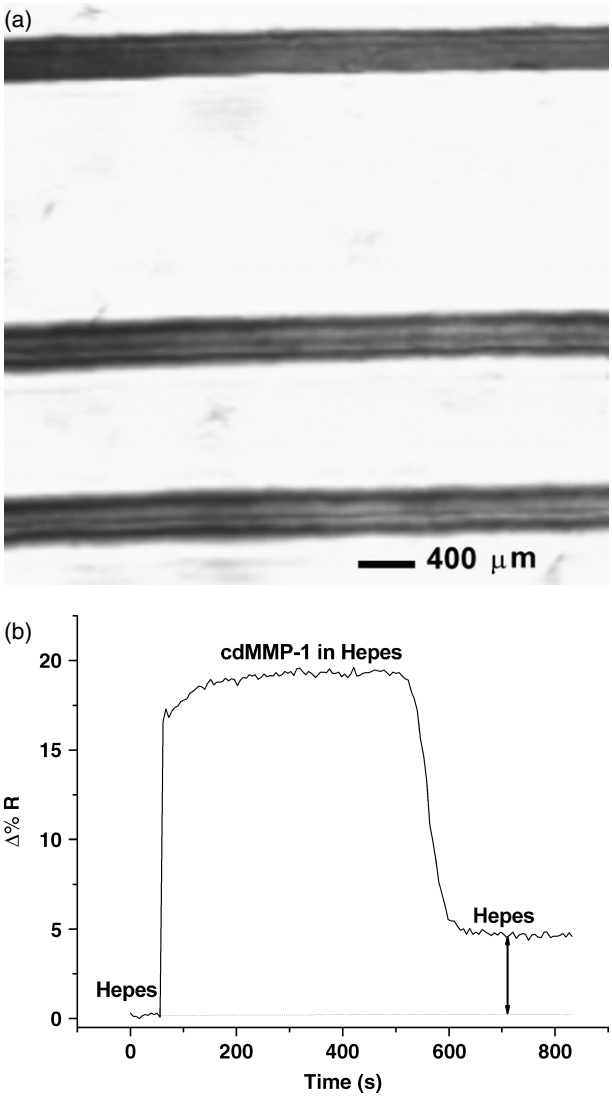


Figure 1. Difference SPR image of a cdMMPs array (a) obtained by flowing cdMMP-1, cdMMP-8 and cdMMP-12 in three different PDMS microchannels (bottom, middle and top, respectively) in contact with a prefunctionalized gold surface as described in the text. (b) The change in percent reflectivity over time obtained for the immobilization of cdMMP-1 reported as an example. The difference between the Δ% R values before and after the protein transit in the microchannel is proportional to the amount of immobilized cdMMP-1.

that a surface coverage of about 4×10^{13} molecules/cm^2 was generally used. For each different protein, 60 μl of *N*-(2-hydroxyethyl)piperazine-*N*′-(2-ethanesulfonic acid) sodium salt (0.01 M HEPES, 0.15 M NaCl, pH 7.4; HBS-N, Biacore International AB) (HBS-N buffer hereafter) was mixed with the appropriate volume of protein mother solution in order to obtain a final concentration of 4.5 μM. This solution obtained was then left in contact with the prefunctionalized[35] gold surface for about 10 min.

In situ protein digestion and AP/MALDI-MS

Structure, stability and orientation of BSA adsorbed to silica has already been investigated in order to draw some conclusions on the structural changes that such protein undergoes when it is adsorbed on a solid support.[44] For

DOI: 10.1002/jms

this reason, although in our case cdMMPs were immobilized on a gold surface, BSA was chosen as a model protein for optimizing our experimental procedure by maximizing the sequence coverage value. Several parameters were considered (protein surface coverage, digestion time, immobilized protein/peptidase ratio, use of $ZipTip_{SCX}$ pipette tips (Millipore), sample/matrix solution ratio, etc.), and in Table 1 the results for the highest sequence coverage values obtained for BSA immobilized on gold are reported. Generally, for the digestion of the anchored enzymes, 20 μl of HCl 10^{-3} M (Sigma–Aldrich) was added to a vial containing 20 μg of either trypsin from porcine pancreas (Sigma–Aldrich, Proteomics grade) or, alternatively, endoproteinase Glu-C from *Staphylococcus aureus* V8. Two microliters of the obtained solution was diluted with 500 μl of NH_4HCO_3 100 mM, pH 7.8 (both Sigma–Aldrich) and the resulting solution was spotted on the protein arrayed surface. The latter was then left in contact with the trypsin solution in a humid chamber at 37 °C for about 6 h (longer digestion times did not improve spectra quality). The final solution (0.5 μl) (hereafter, supernatant) was then collected and spotted on a MALDI plate. Before allowing total evaporation of the drop, 0.5 μl of the matrix solution was applied on top. The best results in terms of spectra quality were achieved by using α-cyano-4-hydroxycinnamic acid (CHCA) in 30% trifluoro acetic acid (TFA) (0.1%) and 70% acetonitrile (C_2H_3N) as the matrix solution at a concentration of 1 mg/ml (all the above material was purchased from Sigma-Aldrich. CHCA was purified by recrystallization from ethanol solution). Any further washing steps of the obtained matrix crystals did not seem to improve the MS spectra quality.[32] The use of $ZipTip_{SCX}$ pipette tips (Millipore) in the spotting procedure helped to improve spectra quality and sequence coverage (Table 2).

All the AP/MALDI-MS measurements were carried out by using a Finnigan LCQ Deca XP PLUS (Thermo Electron Corporation, USA) ion trap spectrometer that was fitted with a MassTech Inc. (USA) AP/MALDI source.[45] The latter consists of a flange containing a computer-controlled X–Y positioning stage and a digital camera, and is powered by a control unit that includes a pulsed nitrogen laser (wavelength 337 nm, pulse width 4 ns, pulse energy 300 μJ, repetition rate up to 10 Hz) and a pulsed dynamic focusing (PDF) module that imposes a delay of 25 μs between the laser pulse and the application of the high voltage to the AP/MALDI target plate. PDF has been shown to improve S/N ratios in the AP-MALDI spectra.[40] Laser power was attenuated to about 55%. The target plate voltage was 1.8 kV. The ion trap inlet capillary temperature was 200 °C. Capillary and tube lens offset voltages of 30 and 15 V, respectively, were applied. Other mass spectrometer parameters were as follows: multipole 1 offset at −3.75 V, multipole 2 offset at −9.50 V, multipole RF amplitude 400 V, lens at −24.0 V and entrance lens at −88.0 V. Automatic gain control (AGC) was turned off and instead the injection time was set to 220 ms and 5 microscans per scan. About 4 min acquisition per sample was usually performed.

Spectra of the *in situ* trypsin digested cdMMPs were acquired in a data-dependent fashion by first acquiring full MS scans from m/z 200 to 2000, followed by MS/MS scans of the most intense ions of the previous full MS scan. MS/MS scans were acquired using an isolation width of 5 m/z, activation q_z of 0.250, activation time of 30 ms and normalized collision energy (NCE) in the range 30–40% depending on the ion. (NCE is the amplitude of the resonance excitation RF voltage scaled to the precursor mass on the basis of the formula: RF amplitude = (NCE%/30%) (precursor ion mass × tick amp slope + tick amp intercept), where the tick amp slope and tick amp intercept are instrument-specific values. For our LCQ Deca, 35% NCE for m/z 1000 = 1.8 V.)

Spectral assignment was obtained by comparing the experimental peaks with the theoretical ones obtained by introducing the three different sequences for cdMMP-12,[46] cdMMP-8 (UniProtKB/Swiss-Prot Accession Number: P22894 Met100-Gly262) and cdMMP-1 (UniProtKB/Swiss-Prot Accession Number: P03956, Val101-Pro269) into the MS-Digest protein digestion simulator (http://prospector.-

Table 1. Experimental peaks of digested BSA previously immobilized on a gold surface reported together with the expected peaks and the sequence coverage value obtained for this model protein. Experimental parameters were: protein surface coverage ≈4 × 10^{13} molecules/cm^2; digestion time = 6 h; trypsin/BSA ratio ≈1/5; use of ZipTip; matrix solution: CHCA

m/z submitted	MH^+ matched	Delta Da	Missed cleavages	Sequence	% Coverage
634.1	634.3789	−0.28	1	(R) GVFRR (D)	23.0
656.3	656.3480	−0.048	1	(R) RDTHK (S)	–
712.7	712.3742	0.33	0	(K) SEIAHR (F)	–
927.6	927.4940	0.11	0	(K) YLYEIAR (R)	–
1163.6	1163.6312	−0.031	0	(K) LVNELTEFAK (T)	–
1283.7	1283.7112	−0.011	0	(R) HPEYAVSVLLR (L)	–
1305.6	1305.7167	−0.12	0	(K) HLVDEPQNLIK (Q)	–
1479.7	1479.7960	−0.096	0	(K) LGEYGFQNALIVR (Y)	–
1504.6	1504.9215	−0.32	1	(K) QTALVELLKHKPK (A)	–
1567.7	1567.7433	−0.043	0	(K) DAFLGSFLYEYSR (R)	–
1639.9	1639.9383	−0.038	1	(R) KVPQVSTPTLVEVSR (S)	–
1890.7	1890.8036	−0.10	1	(K) VASLRETYGDMADCCEK (Q)	–
1955.9	1955.9602	−0.060	0	(K) DAIPENLPPLTADFAEDK (D)	–

DOI: 10.1002/jms

Table 2. Experimental peaks observed after digestion with trypsin of cdMMP-8. There are no differences between the spectra obtained from the solution or the supernatant liquid, showing that the immobilization of the enzyme does not affect the peptidase access to the enzyme. Furthermore, the presence of a strong inhibitor (NNGH) does not have a large impact on the sequence coverage. On the contrary, the latter is improved considerably by the use of $ZipTip_{SCX}$ pipette tips. Similar results were found also for cdMMP-1 and cdMMP-12

Expected fragments	Supernatant ZipTip	Supernatant, no ZipTip	Solution ZipTip	Solution, no ZipTip	Solution + NNGH ZipTip	Supernatant + NNGH ZipTip
490.2	√	√	√	√	√	√
767.4	√	√	√	√	√	√
857.5	√	–	√	–	√	–
873.5	–	–	–	–	–	–
1036.6	–	–	–	–	–	–
1238.6	√	–	√	–	√	–
1328.7	√	–	√	–	√	–
1344.7	–	–	–	–	–	–
1507.8	–	–	–	–	–	–
1535.7	√	√	√	√	√	√
1724.9	√	√	√	√	√	√
1804.9	–	–	–	–	–	–
1847.9	–	–	–	–	–	–
1852.0	√	√	√	√	–	√
Coverage %	37.4	32.5	37.4	32.5	27.6	32.5

ucsf.edu/ucsfhtml4.0/msdigest.htm). In the case of cdMMP-12 and cdMMP-1, ^{15}N enrichment was also considered.

RESULTS AND DISCUSSION

The immobilization of cdMMPs on solid substrates represents a fundamental step for the development of solid-state MMP assays. It has already been highlighted[35] how an improper surface packing density of cdMMPs facilitates the autodegradation of the enzyme and affects the activity of the immobilized enzyme itself. At the same time, a higher protein density at the solid surface is desirable for analytical purposes owing to the resulting increase in the MS signals of the species formed after the *in situ* tryptic digestion. For these reasons, AP/MALDI-MS experiments have been carried out by using gold substrates on which cdMMP-1, cdMMP-8 and cdMMP-12 were immobilized at the maximum surface density that preserved the enzyme activity. Both the cdMMPs surface coverage and the immobilized cdMMPs activity have been evaluated by following the already described procedures.[35] In particular, for all the experiments discussed below, a surface coverage of about 4×10^{13} molecules/cm^2 was used.

The use of MALDI-MS for detecting chemical processes that involve MMP molecules is not something completely new.[47] Nevertheless, in almost all the works available in the current literature,[48] it is common practice to detect only the products of MMP proteolytic activity by MALDI-MS, while the tryptic digestion of the enzyme itself has never been reported.[49,50] This is probably due to the difficulties in obtaining detectable amounts of purified enzyme not being mixed with buffer and salt-contaminated solutions, as it is known that MMPs require a buffer containing specific amounts of zinc and calcium cations, as well as proper pH values, in order to maintain their structural integrity and biological activity.[51,52] Even for our stabilized forms of cdMMP-1, cdMMP-8 and cdMMP-12,[46] the starting solutions, respectively 90 μM (**1**), 50 μM (**2**) and 270 μM (**3**) in enzyme, were also 20 nM in the Tris buffer, 10 mM in $CaCl_2$, 0.1 mM in $ZnCl_2$ and 200 mM in AHA. It is then clear that the amount of salts added in order to have a stabilized and active form of the enzymes is larger than the enzymes themselves. This can cause problems in the MALDI-MS spectrum in terms of both sensitivity and peak detection.

In Fig. 1(a), the SPR image of a gold surface arrayed with the three different cdMMPs is reported, together with a representative graph showing the change in percent reflectivity over time obtained for the cdMMP-1 immobilization (Fig. 1(b)). Appropriate separate spotting of the trypsin solution on each different channel (Fig. 2) allowed the spatially resolved recognition of the three proteins, producing the MS spectra in the m/z 200–2000 range reported in Fig. 3 (the MS spectrum for the same samples in the range 2000–4000 are not shown as they are not informative for our purpose). It is possible to see that some peaks are detected in the sodiated form[34] (m/z 865.6, 1271.7 and 1591.7 in Fig. 3(a), m/z 1260.6 in Fig. 3(b)) and, in some cases, the peak assignments had to be confirmed by MS/MS experiments in order to avoid ambiguity (Fig. 4). In the case of cdMMP-8 (Fig. 3(b)), several other unassigned peaks are also observed and this is due to the fact that the problem of auto-hydrolysis, commonly encountered for all metalloproteinases,[53] is more severe for this member of the MMP family.

The presence of contaminants is easily revealed by performing MS analyses of blank areas treated as above but without the presence of cdMMPs. In this case several peaks originating from matrix clusters with Na^+ are observed, as

DOI: 10.1002/jms

JMS

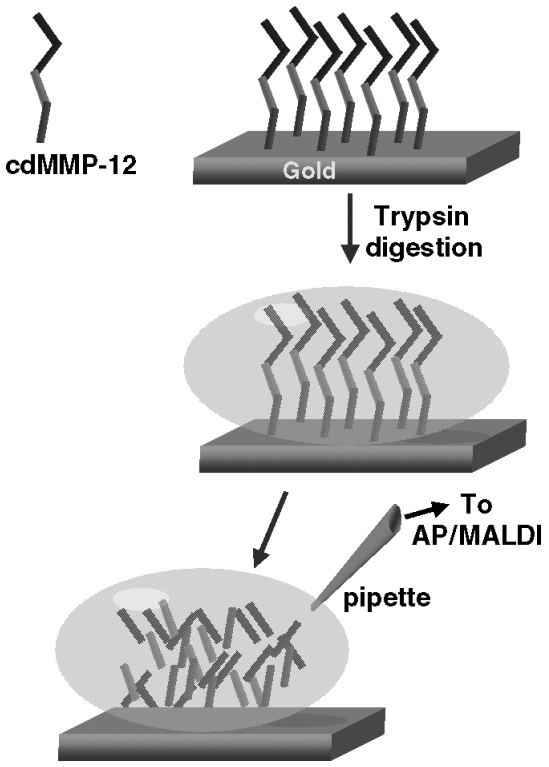


Figure 2. Scheme of the procedure applied for *in situ* tryptic digestion and AP/MALDI-MS characterization. Enzyme molecules previously immobilized on the gold surface (cdMMP-12 in this particular case) are put in contact with the tryptic solution in a humid chamber at 37 °C overnight. The supernatant solution (0.5 μl) is sampled after 6 h for AP/MALDI-MS analysis. The same experimental procedure was applied also using a different peptidase, endoproteinase Glu-C, and the results are discussed in the text.

shown in Fig. 5. In our case, the use of ZipTip helped to remove some of the contaminants and it gave an important contribution for the acquisition of the less salt-affected MS spectra.

It has been reported that when immobilized to a solid support, some proteins are less accessible to the action of peptidases, producing a smaller sequence coverage.[44] In most cases, by reducing the degrees of freedom, both activity and structure of the immobilized protein are usually affected and the general trend is that structurally less stable proteins lose more native structure upon adsorption than more stable proteins.[54] Nevertheless, the impact that adsorption to a solid support has on the properties of a protein depends greatly on the nature of such adsorption. Usually, in order to study biological molecules by SPR, a low-impact adsorption is highly desirable. Ideally, it should be possible to array proteins on a surface and the anchored biomolecules should preserve their properties as in the solution state, so that interaction with other molecules is meaningful. We have already proved that anchored cdMMP-1 has the same activity as in the solution state,[35] and here the immobilization procedure is expanded to two other cdMMPs and it is shown that it does not affect the action of peptidases for protein degradation and recognition by peptide fingerprint. We also inhibited

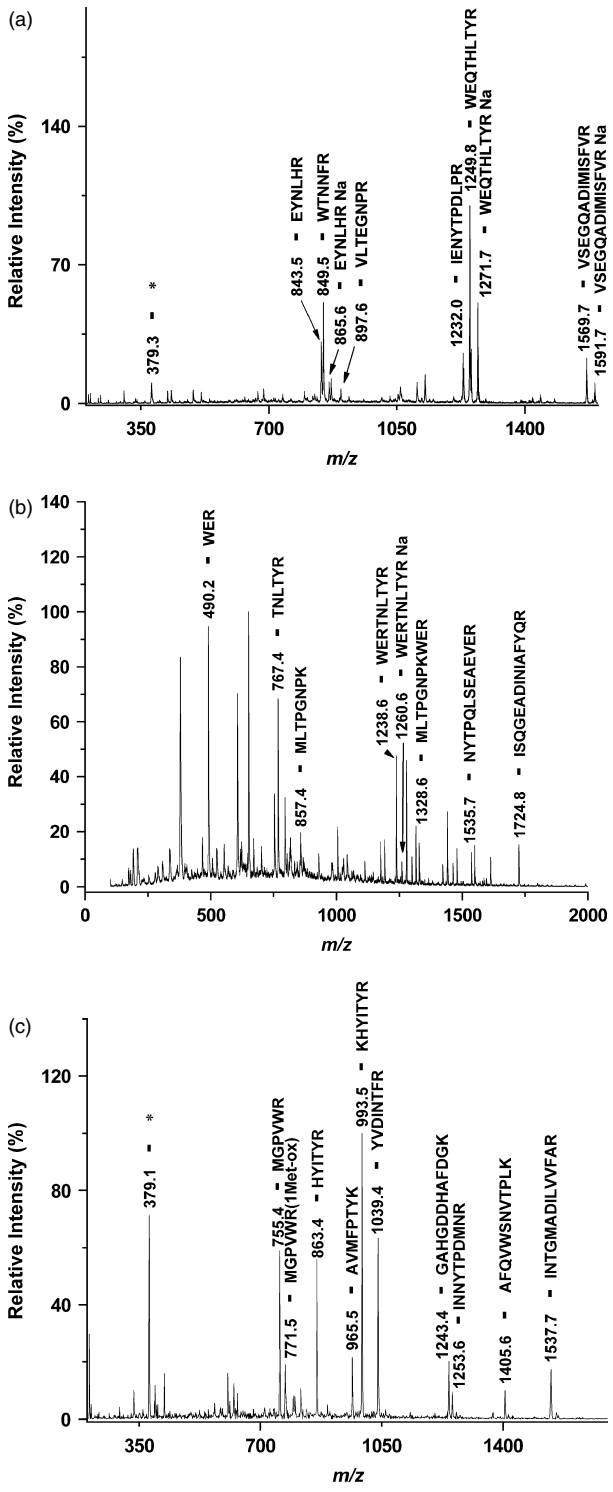


Figure 3. AP/MALDI-MS spectra of three different proteins, cdMMP-1 (a), cdMMP-8 (b) and cdMMP-12 (c), after immobilization on a gold substrate and on-chip digestion with trypsin solution. The assignment of the most intense peaks is reported. Peaks attributed to the matrix crystals are labelled with *, and the positive charge is omitted for simplicity. Some of the peaks (*m/z* 865.6, 1271.7 and 1591.7 in (a), *m/z* 1260.6 in (b)) were assigned to sodiated peptides that form because of the high salt content of the initial protein solution. Although many unassigned peaks are present in (b) owing to protein auto-hydrolysis, even in this case the sequence coverage has an acceptable value, which ensures unambiguous characterization of the protein (Table 3).

DOI: 10.1002/jms

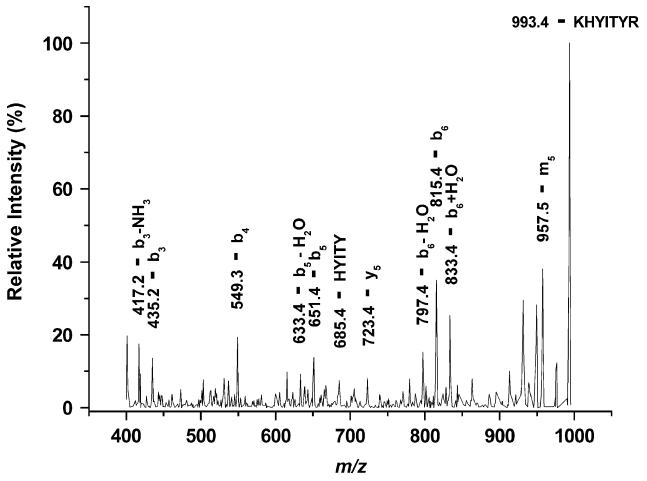


Figure 4. MS/MS fragmentation pattern for the *m/z* 993.4 peak of the cdMMP-12 spectrum. Because of the ^{15}N enrichment, the peak is assigned to the fragment KHYITYR having *m/z* 980.5 without the ^{15}N enrichment. The observed fragmentation pattern confirmed our assignment, excluding the possible presence of the fragment EDVDYAIR, having *m/z* 980.4 without the ^{15}N enrichment, but a different fragmentation pattern.

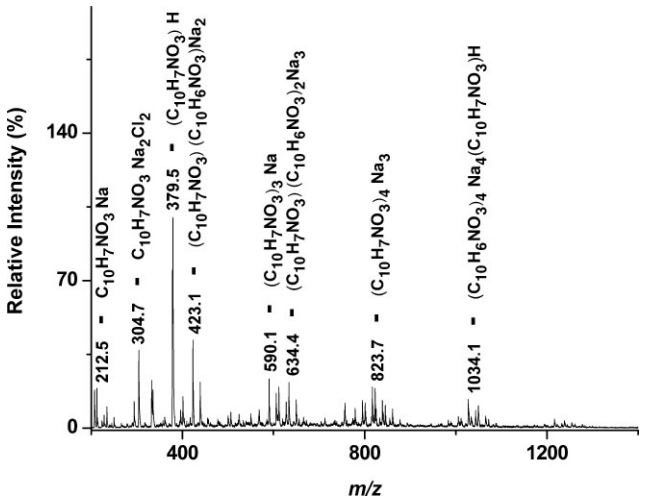


Figure 5. AP/MALDI-MS spectrum obtained by sampling the supernatant solution from an area of the gold surface where no enzyme was immobilized. Many peaks that could be assigned to CHCA clusters containing Na^+ ions are visible, confirming that a salt-free MS spectrum can be a challenging task in such experimental conditions. In fact, the presence of buffered solutions is necessary to maintain enzyme activity, as discussed in the text.

the cdMMPs with a strong inhibitor, *N*-isobutyl-*N*-(4-methoxyphenylsulfonyl)-glycylhydroxamic acid (NNGH), and only negligible differences were observed in terms of peptide appearance in the mass spectra of the immobilized enzyme (Table 2). The use of Glu-C as peptidase lowered the sequence coverage values for all of the cdMMPs investigated, so trypsin was selected as optimal peptidase in these conditions (Table 3). Generally, the high sequence coverage values obtained when trypsin was used confirmed the idea that our immobilization procedure does not have any MS-measurable impact on protein properties.[35] This is probably due to the fact that, by using amine coupling chemistry, anchored proteins are randomly oriented on the surface (interacting $-NH_2$ groups in cdMMPs are equally distributed on protein surfaces), so that, overall, all the sides of the biomolecules are equally exposed to the action of peptidases.

In order to extend the applicability of our multiplexed approach to real biological samples,[55] we applied the same experimental procedure to a 1 : 1 mixture of cdMMP-1/cdMMP-12 immobilized on the solid surface. From the results shown in Fig. 6, it is possible to see that almost all the MS peaks detected for single protein solutions are found also in this case. The same result was obtained also for cdMMP-8/cdMMP-12 mixture (data not shown) and this confirms that the experimental protocol is applicable to protein mixtures. The same MS spectra were recorded with the

Table 3. Sequence coverage values for the digestion of cdMMP-1, cdMMP-8, cd-MMP-12 and BSA with two different peptidases, trypsin and endoproteinase Glu-C. The latter gives usually lower sequence coverage values in the experimental conditions used, so that trypsin was chosen as the optimal enzyme for our experimental protocol

Protein	Sequence coverage (%)	
	Trypsin	Endoproteinase Glu-C
cdMMP-1	31.1	20.5
cdMMP-8	37.4	31.9
cdMMP-12	47.7	25.5
BSA	30.6	20.6

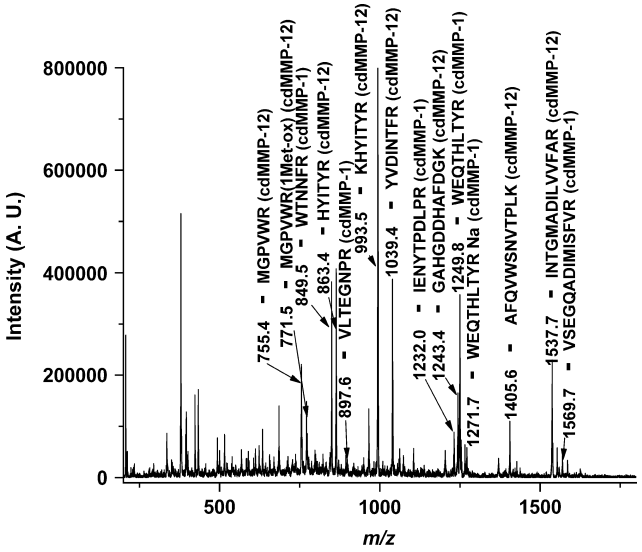


Figure 6. AP/MALDI-MS spectrum obtained by sampling the supernatant solution from an area of the gold surface where both cdMMP-1 and cdMMP-12 were previously immobilized and digested with trypsin solution as described in the text. The same MS spectrum was recorded with the protein molecules mixed before or after the immobilization on the gold surface. Although the spectrum is more complicated because it is crowded with peaks derived from the tryptic digestion of both cdMMPs, a meaningful assignment is also possible in this case, confirming the applicability of the experimental procedure to protein mixtures.

DOI: 10.1002/jms

protein molecules mixed before or after the immobilization on the gold surface, proving once again that the anchorage on the solid support does not alter the protein interaction properties.

From all the results shown above, we therefore think that the new experimental procedure is particularly suitable for SPR/MALDI-MS coupled investigation of enzyme activity and/or interactions with other molecules, and it can be developed further in order to extend its applicability to real complex biological samples.

CONCLUSIONS

The obtained results for cdMMP-1, cdMMP-8 and cdMMP-12 show that it is possible to characterize an enzyme that had been previously anchored on a gold surface by *in situ* AP/MALDI-MS. Although such experimental protocol can have many different applications, we believe that the new approach could be particularly advantageous for the development of solid-state MMP assays. Moreover, it could be of help to SPR users, as the necessity to unambiguously characterize biomolecules that have been previously studied by SPR is often an important issue in the BIA field. Finally, the obtained results are a further proof of the importance that buffer- and salt-contaminated solution has in the MS analysis, not only in diminishing the sensitivity but also in affecting the appearance of the MS spectrum.

The applicability of the above-described experimental approach to more complicated samples (mixture of cdMMPs), as well as to protein arrays, has also been demonstrated, proving that the experimental protocol can be generally used for unambiguously identifying immobilized proteins on solid supports.

Acknowledgements

We thank MIUR (FIRB 2003 RBNE01TTJW_004 and PRIN 2005 n 2 005 038 704_004) for partial financial support.

REFERENCES

1. MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. *Science* 2000; **289**: 1760.
2. Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, Lan N, Jansen R, Bidlingmaier S, Houfek T, Mitchell T, Miller P, Dean RA, Gerstein M, Snyder M. Global analysis of protein activities using proteome chips. *Science* 2001; **293**: 2101.
3. Houseman BT, Huk JH, Kron SJ, Mrksich M. Peptide chips for the quantitative evaluation of protein kinase activity. *Nat. Biotechnol.* 2002; **20**: 270.
4. Huber A, Demartis S, Neri D. The use of biosensor technology for the engineering of antibodies and enzymes. *J. Mol. Recognit.* 1999; **12**: 198.
5. Reymond JL. *Enzyme Assays: High-throughput Screening, Genetic Selection and Fingerprinting*. Wiley-VCH Verlag GmbH and Co. KGaA: Weinheim, 2006.
6. Min D, Mrksich M. Profiling kinase activities by using a peptide chip and mass spectrometry. *Angew. Chem., Int. Ed. Engl.* 2004; **43**: 5973.
7. Babiak P, Reymond J. A high-throughput, low-volume enzyme assay on solid support. *Anal. Chem.* 2005; **77**: 373.
8. Huber W, Hoffman F. A new strategy for improved secondary screening and lead optimization using high-resolution SPR characterization of compound-target interactions. *J. Mol. Recognit.* 2005; **18**: 273.
9. Nedelkov D, Nelson RW. Surface plasmon resonance mass spectrometry for protein analysis. *Methods Mol. Biol.* 2006; **328**: 131.
10. Nedelkov D, Nelson RW. Surface plasmon resonance mass spectrometry: recent progress and outlooks. *Trends Biotechnol.* 2003; **21**: 301.
11. Borch J, Roepstorff P. Screening for enzyme inhibitors by surface plasmon resonance combined with mass spectrometry. *Anal. Chem.* 2004; **76**: 5243.
12. Nedelkov D, Nelson RW. Design and use of multi-affinity surfaces in biomolecular interaction analysis-mass spectrometry (BIA/MS): a step toward the design of SPR/MS arrays. *J. Mol. Recognit.* 2003; **16**: 15.
13. Borch J, Roepstorff P. Combinations of SPR and MS for characterization of native and recombinant proteins in cell lysates. *Mol. Biotechnol.* 2006; **33**: 179.
14. Hutchens T, Yip TT. New desorption strategies for the mass spectrometric analysis of macromolecules. *Rapid Commun. Mass Spectrom.* 1993; **7**: 576.
15. Caputo E, Moharram R, Martin BM. Methods for on-chip analysis. *Anal. Biochem.* 2003; **321**: 116.
16. Merchant M, Weinberger SR. Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. *Electrophoresis* 2000; **21**: 1164.
17. Griesser HJ, Kingshott P, McArthur SL, McLean KM, Kinsel GR, Timmons RB. Surface-MALDI mass spectrometry in biomaterials research. *Biomaterials* 2004; **25**: 4861.
18. Clarke CH, Buckley JA, Fung ET. SELDI-TOF-MS proteomics of breast cancer. *Clin. Chem. Lab. Med.* 2005; **43**: 1314.
19. Bons JAP, Wodzig WKWH, van Dieijen-Visser MP. Protein profiling as a diagnostic tool in clinical chemistry: a review. *Clin. Chem. Lab. Med.* 2005; **43**: 1281.
20. Issaq HJ, Conrads TP, Prieto DA, Tirumalai R, Veenstra TD. SELDI-TOF MS for diagnostic proteomics. *Anal. Chem.* 2003; **1**: 149A.
21. Ge Y, Gibbs BF, Masse R. Complete chemical and enzymatic treatment of phosphorylated and glycosylated proteins on proteinchip arrays. *Anal. Chem.* 2005; **77**: 3644.
22. Shen Z, Go EP, Gamez A, Apon JV, Fokin V, Greig M, Ventura M, Crowell JE, Blixt O, Paulson JC, Stevens RC, Finn MG, Siuzdak G. A mass spectrometry plate reader: monitoring enzyme activity and inhibition with a Desorption/Ionization on Silicon (DIOS) platform. *Chembiochem* 2004; **5**: 921.
23. Min D, Tang W, Mrksich M. Chemical screening by mass spectrometry to identify inhibitors of anthrax lethal factor. *Nat. Biotechnol.* 2004; **22**: 717.
24. Bornsen KO, Gass MAS, Bruin GJM, von Adrichem JHM, Biro MC, Kresbach GM, Ehrat M. Influence of solvents and detergents on matrix-assisted laser desorption/ionization mass spectrometry measurements of proteins and oligonucleotides. *Rapid Commun. Mass Spectrom.* 1997; **11**: 603.
25. Brockman AH, Dodd BS, Orlando N. A desalting approach for MALDI-MS using on-probe hydrofobic self-assembled monolayers. *Anal. Chem.* 1997; **69**: 4716.
26. Nedelkov D, Nelson RW. Practical considerations in BIA/MS: optimizing the biosensor-mass spectrometry interface. *J. Mol. Recognit.* 2000; **13**: 140.
27. Worral TA, Cotter RJ, Woods AS. Purification of contaminated peptides on synthetic membrane surfaces for matrix-assisted laser desorption/ionization mass spectrometry. *Anal. Chem.* 1998; **70**: 750.
28. Warren ME, Brockman AH, Orlando R. On-probe solid-phase extraction/MALDI-MS using ion-pairing interactions for the cleanup of peptides and proteins. *Anal. Chem.* 1998; **70**: 3757.
29. Xu Y, Watson JT, Bruening ML. Patterned monolayer/polymer films for analysis of dilute or salt-contaminated protein samples by MALDI-MS. *Anal. Chem.* 2003; **75**: 185.
30. Wang Y, Xia X, Guo Y. Porous anodic alumina membrane as a sample support for MALDI-TOF MS analysis of salt-containing proteins. *J. Am. Soc. Mass Spectrom.* 2005; **16**: 1488.

31. Smirnov IP, Zhu X, Taylor T, Huang Y, Ross P, Papayanopoulos IA, Martin SA, Pappin DJ. Suppression of α-Cyano-4-hydroxycinnamic acid matrix clusters and reduction of chemical noise in MALDI-TOF mass spectrometry. *Anal. Chem.* 2004; **76**: 2958.
32. Kim J-S, Kim J-Y, Kim H-J. Suppression of matrix clusters and enhancement of peptide signals in MALDI-TOF mass spectrometry using nitrilotriacetic acid. *Anal. Chem.* 2005; **77**: 7483.
33. Newton KA, McLuckey SA. Generation and manipulation of Sodium cationized peptides in the gas phase. *J. Am. Soc. Mass Spectrom.* 2004; **15**: 607.
34. Su J, Mrksich M. Using MALDI-TOF mass spectrometry to characterize interfacial reactions on self-assembled monolayers. *Langmuir* 2003; **19**: 4867.
35. Grasso G, D'Agata R, Rizzarelli E, Spoto G, D'Andrea L, Pedone C, Picardi A, Romanelli A, Fragai M, Yeo KJ. Activity of anchored human matrix metalloproteinase-1 catalytic domain on au (111) surfaces monitored by ESI-MS. *J. Mass Spectrom.* 2005; **40**: 1565.
36. Massova I, Kotra LP, Fridman R, Mobashery S. Matrix metalloproteinases: structures, evolution, and diversification. *FASEB J.* 1998; **12**: 1075.
37. Coussens LM, Werb Z. Matrix metalloproteinases and the development of cancer. *Chem. Biol.* 1996; **3**: 895.
38. Koivunen E, Arap W, Valtanen H, Rainisalo A, Medina OP, Heikkilä P, Kantor C, Gahmberg CG, Salo T, Konttinen YT, Sorsa T, Ruoslahti E, Pasqualini R. Tumor targeting with a selective gelatinase inhibitor. *Nat. Biotechnol.* 1999; **17**: 768.
39. Farkas E, Katz Y, Bhusare S, Reich R, Röschenthaler G, Königsmann M, Breuer E. Carbamoylphosphonate Based Matrix Metalloproteinase (MMP) inhibitor metal complexes–solution studies and stability constants. Towards a zinc selective binding group. *J. Biol. Inorg. Chem.* 2004; **9**: 307.
40. Schneider BB, Lock C, Covey TR. AP and vacuum MALDI on a QqLIT instrument. *J. Am. Soc. Mass Spectrom.* 2005; **16**: 176.
41. Trimpin S, Deinzer ML. Solvent-free MALDI-MS for the analysis of biological samples via a mini-ball mill approach. *J. Am. Soc. Mass Spectrom.* 2005; **16**: 542.
42. Whitesides G, Xia Y. Soft lithography. *Angew. Chem., Int. Ed. Engl.* 1998; **37**: 550.
43. Navratilova I, Eisenstien E, Myszka DG. Measuring long association phases using Biacore. *Anal. Biochem.* 2005; **344**: 295.
44. Larsericsdotter H, Oscarsson S, Buijs J. Structure, stability, and orientation of BSA adsorbed to silica. *J. Colloid Interface Sci.* 2005; **289**: 26.
45. Tan PV, Laiko VV, Doroshenko VM. Atmospheric pressure MALDI with pulsed dynamic focusing for high-efficiency transmission of ions into a mass spectrometer. *Anal. Chem.* 2004; **75**: 2462.
46. Bertini I, Calderone V, Cosenza M, Fragai M, Lee Y-M, Luchinat C, Mangani S, Terni B, Turano P. Conformational variability of matrix metalloproteinases: beyond a single 3D structure. *Proc. Natl. Acad. Sci. U.S.A.* 2005; **102**: 5334.
47. Buchardt J, Schiodt CB, Krog-Jensen C, Delaisse J-M, Foged NT, Meldal M. Solid phase combinatorial library of phosphinic peptides for discovery of matrix metalloproteinase inhibitors. *J. Comb. Chem.* 2000; **2**: 624.
48. Feng R, Castelhano AL, Billedeau R, Yuan Z. Study of noncovalent enzyme-inhibitor complexes and metal binding stoichiometry of matrilysin by electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* 1995; **6**: 1105.
49. Sung JY, Park SM, Lee C-H, Um JW, Lee HJ, Kim J, Oh YJ, Lee S-T, Paik SR, Chung KC. Proteolytic cleavage of extracellular secreted a-synuclein via matrix metalloproteinases. *J. Biol. Chem.* 2005; **280**: 25 216.
50. Lauer-Fields JL, Nagase H, Fields GB. Use of edman degradation sequence analysis and matrix-assisted laser desorption/ionization mass spectrometry in designing substrates for matrix metalloproteinases. *J. Chromatogr., A* 2000; **890**: 117.
51. Gossas T, Danielason UH. Characterization of Ca^{2+} interactions with matrix metallopeptidase-12: implications for matrix metallopeptidase regulation. *Biochem. J.* 2006; **398**: 393.
52. Willenbrock F, Murphy G, Phillips IR, Brocklehurst K. The second zinc atom in the matrix metalloproteinase catalytic domain is absent in the full-length enzymes: a possible role for the C-terminal domain. *FEBS Lett.* 1995; **358**: 189.
53. Knaeuper V, Docherty AJP, Smith B, Tschesche H, Murphy G. Analysis of the contribution of the hinge region of human neutrophil collagenase (HNC, MMP-8) to stability and collagenolytic activity by alanine scanning mutagenesis. *FEBS Lett.* 1997; **405**: 60.
54. Kondo A, Urabe T. Temperature dependence of activity and conformational changes in a-amylases with different thermostability upon adsorption on ultrafine silica particles. *J. Colloid Interface Sci.* 1995; **174**: 191.
55. Keller A, Purvine S, Nesvizhskii A, Stolyar S, Goodlett DR, Kolker E. Experimental protein mixture for validating tandem mass spectral analysis. *OMICS* 2002; **6**: 207.

DOI: 10.1002/jms

4.4 Snapshots of the Reaction Mechanism of Matrix Metalloproteinases [*Angew. Chem. Int. Ed. Engl.*, 2006, 45, 7952-7955]

Enzyme Mechanisms

DOI: 10.1002/anie.200603100

Snapshots of the Reaction Mechanism of Matrix Metalloproteinases**

Ivano Bertini, Vito Calderone, Marco Fragai, Claudio Luchinat, Massimiliano Maletta, and Kwon Joo Yeo*

Matrix metalloproteinases (MMPs) are a family of proteins involved in tissue remodeling and cell signaling.[1–3] As such, they are validated drug targets.[4,5] MMPs are endopeptidases with zinc at the active site of a catalytic domain.[6]

The enzymology of peptidases has been a popular area of research since the investigation of carboxypeptidase A (CPA)[7–11] and thermolysin (TLN)[12–15] in the 1980s. Both proteins are rich in charged and ionizable groups, such as Glu, Arg, and His, in the active site. A role for these residues in catalysis was proposed. In contrast, MMPs are simple proteins with only one Glu residue in the active site besides the three His residues coordinated to the catalytic zinc ion. In the most widely accepted mechanism, which is similar to that proposed for CPA and TLN, this Glu residue assists in the nucleophilic attack at the relevant peptide carbonyl group of the substrate by a zinc-coordinated water molecule.[16] The interest in MMPs has grown again because of the need for inhibitors that are as specific as possible for each of the 23 human proteins.[17–22]

We have now solved a series of X-ray crystal structures under a variety of conditions in an attempt to obtain models of the various steps of the reaction mechanism of MMPs. The protein tends to digest itself, so crystallization is always carried out in the presence of an inhibitor. By experimenting with various crystals under crystal-washing conditions we could solve the structure of the uninhibited form of MMP-12 at a resolution of 1.2 Å (Figure 1a; see also the Supporting Information) and that of MMP-8 at a resolution of 1.7 Å (see the Supporting Information). These are the first X-ray crystal structures of the uninhibited, active form of MMP. In MMP-12, the active site contains three water molecules coordinated to zinc, one of which is hydrogen bonded to Glu219, in an almost regular octahedral geometry. In the case of MMP-8, only the water molecule that is hydrogen bonded to Glu219 has full occupancy, whereas the other two water molecules have occupancies of only about 20%. Despite the lower resolution of the MMP-8 structure, the electron densities of the latter water molecules are clearly appreciable (see the Supporting Information). If these two water molecules are neglected, the resulting geometry is a flattened tetrahedron, with ample room for the binding of an exogenous ligand.

We then attempted to soak the 204–209 peptide fragment of type I/III of the alpha-1 collagen chain, *ProGlnGlyIle-AlaGly*, which is known to be cleaved at the *Gly–Ile* bond, into the active MMP-12 and MMP-8 crystals. No peptide-bound form could be identified, as a result of rapid hydrolysis. Even with the mutation Glu219Ala, which had been reported to lower the enzymatic activity of MMP-7 by three orders of magnitude,[23] the substrate underwent hydrolysis. On the other hand, well-resolved X-ray crystallographic structures that show a hydrolysis product inside the active site could be obtained for both MMP-12 (resolution: 1.1 Å; Figure 1d) and MMP-8 (resolution: 1.5 Å; see the Supporting Information), and for the former even a structure with both hydrolysis products was obtained (resolution: 1.8 Å; Figure 1c). We then modeled the substrate in the form of a tetrahedral *gem*-diol intermediate (in analogy to the structure of a pseudosubstrate described in the literature[24]) by taking the uninhibited form of MMP-12 and that with two peptides in the active site as the starting points (Figure 1b). Figure 1a–d show a plausible series of events in the catalytic cycle of MMPs, as discussed below. A four-snapshot movie of the steps illustrated in Figure 1a–d can be downloaded as Supporting Information for this article.

The cycle starts with the uninhibited MMP-12 enzyme (Figure 1a). Figure 1b shows that one of the two oxygen atoms of the *gem*-diol that results from substrate binding

[*] Prof. I. Bertini, V. Calderone, M. Fragai, C. Luchinat, M. Maletta, K. J. Yeo
Magnetic Resonance Center (CERM)
University of Florence
Via L. Sacconi 6, 50019 Sesto Fiorentino (Italy)
Fax: (+39) 055-457-4271
E-mail: ivanobertini@cerm.unifi.it
Prof. I. Bertini, M. Maletta
Department of Chemistry
University of Florence
Via della Lastruccia 3, 50019 Sesto Fiorentino (Italy)
M. Fragai, C. Luchinat
Department of Agricultural Biotechnology
University of Florence
Via Maragliano 75–77, 50144 Florence (Italy)

[**] We thank Dr. Andrea Giachetti for help with computational analysis. This research was supported by the EC (projects: LSHG-CT-2004-512077 and "Structural Proteomics in Europe" Grant QLG2-CT-2002-00988), the MIUR (PRIN 2005, Prot. N. 2005039878, Prot. N. 4455, and RBAU013NSB), and Ente Cassa di Risparmio di Firenze (Innovative Strategies for Drug Design). We acknowledge the support and assistance of the ESRF (Grenoble) and DESY (Hamburg) synchrotron radiation facilities for the data collection.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

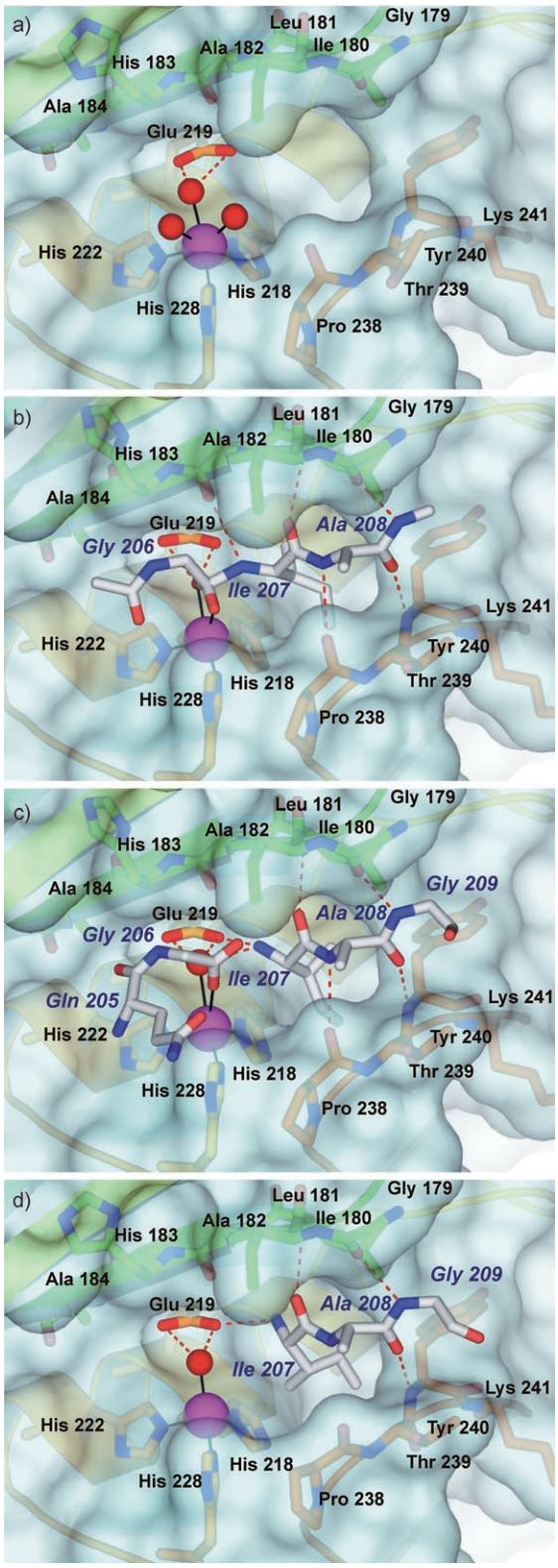


Figure 1. a) Active, uninhibited form of MMP-12. b) A plausible *gem*-diol intermediate[24] of MMP-12 modeled by using the structures of the uninhibited (a) and two-peptide (c) forms as templates. c) Two-peptide intermediate observed upon soaking the active uninhibited MMP-12 crystals with the collagen fragment *ProGlnGlyIleAlaGly*. d) *IleAlaGly* adduct of MMP-12.

indeed sits in the position previously occupied by the catalytically relevant water molecule and is also hydrogen bonded to Glu 219, whereas the other occupies a coordination site between the positions of the other two more external water molecules in Figure 1 a. The peptide NH group of the scissile bond is hydrogen bonded to the carbonyl oxygen atom of Ala 182, and the side chain of *Ile 207* partially enters the S_1' cavity, whereas the rest of the body of the peptide is stabilized by four hydrogen-bonding interactions with the backbone of the protein. The latter interactions involve the carbonyl oxygen atom of *Ile 207* and the nitrogen atom of Leu 181, the nitrogen atom of *Ala 208* and the carbonyl oxygen atom of Pro 238, the carbonyl oxygen atom of *Ala 208* and the nitrogen atom of Tyr 240, and the nitrogen atom of *Gly 209* and the carbonyl oxygen atom of Gly 179. As the modeling did not provide a convincing conformation for the Pro and Gln residues of the peptide, they are omitted from Figure 1 b.

After peptide-bond cleavage, both hydrolysis fragments, *ProGlnGly* and *IleAlaGly*, sit in the active-site cavity, as shown in Figure 1 c. The peptide *ProGlnGly*, for which a clear electron density is observed only for the whole *Gly 206* residue and part of *Gln 205*, is apparently not involved in any significant stabilizing interaction with the protein backbone. The carboxylate end of *Gly 206* binds the zinc ion in an *anti*-monodentate fashion. A water molecule is also semicoordinated to zinc (Zn–O distance: 2.8 Å) and hydrogen bonded to the catalytically relevant Glu 219 residue; it acts as a "spacer" between Glu 219 and the coordinated carboxylate end of *Gly 206*. The zinc ion is thus five-coordinated in this intermediate. The semicoordinated water molecule occupies a position similar to that of the catalytically relevant water molecule observed in the free uninhibited enzyme. The full *IleAlaGly* fragment can be detected clearly. Its orientation in the active site is such that the side chain of *Ile 207* partially enters the S_1' cavity, whereas the rest of the body of the peptide is stabilized by four hydrogen-bonding interactions with the backbone of the protein. The hydrogen-bonding distances are 2.91 Å between the carbonyl oxygen atom of *Ile 207* and the nitrogen atom of Leu 181, 3.16 Å between the nitrogen atom of *Ala 208* and the carbonyl oxygen atom of Pro 238, 2.81 Å between the carbonyl oxygen atom of *Ala 208* and the nitrogen atom of Tyr240, and 2.71 Å between the nitrogen atom of *Gly 209* and the carbonyl oxygen atom of Gly 179. The N-terminal nitrogen atom of the *IleAlaGly* fragment is involved in two strong hydrogen-bonding interactions with the water molecule coordinated to the zinc atom (2.69 Å) and with one oxygen atom of Glu 219 (2.24 Å), whereas there are no hydrogen-bonding interactions of this nitrogen atom with the protein backbone. It is difficult to ascertain whether the N-terminal nitrogen atom is part of a neutral NH_2 group or a positive NH_3^+ group. Overall, the active-site crevice is slightly more open than in the active uninhibited enzyme, in contrast to what was observed for thermolysin and astacin.[25,26]

The *ProGlnGly* fragment, which is coordinated to zinc through the carboxylate group of the C-terminal *Gly 206* residue, is the first to leave the cavity, whereas the *IleAlaGly* fragment remains in the cavity (Figure 1 d; see also the Supporting Information). The zinc ion appears in this case to

be tetracoordinated, with a more strongly bound water molecule (2.28 Å), which is in turn hydrogen-bonded to Glu219. The interaction between this water molecule and Glu219 is still very strong; they are separated by a distance of 2.60 Å. This water molecule occupies roughly the same position as in the free uninhibited enzyme, but the other two coordinated water molecules are not present, so that the zinc ion has a flattened tetrahedral coordination sphere. Given the high resolution of the structure, the existence of other immobilized water molecules in the zinc coordination sphere can be safely ruled out (see the Supporting Information). The bond lengths around the zinc ion (see the Supporting Information) are also consistent with a lower coordination number of the zinc center. The amino terminus of the *IleAlaGly* fragment moves away from the zinc center and loses the strong hydrogen-bonding interaction with the zinc-coordinated water molecule, whereas the interaction with one of the oxygen atoms of Glu219 (2.74 Å) is retained. The side chain of *Ile207* alters its position and increases its interactions with the S_1' cavity. The rest of the body of the peptide is stabilized by three of the four hydrogen-bonding interactions with the backbone of the protein that were observed when two peptides occupied the active site (carbonyl oxygen atom of *Ile207* with the nitrogen atom of Leu181, 2.77 Å; carbonyl oxygen atom of *Ala208* with the nitrogen atom of Tyr240, 2.74 Å; nitrogen atom of *Gly209* with the carbonyl oxygen atom of Gly179, 3.03 Å). However, the interaction between the nitrogen atom of *Ala208* and the carbonyl oxygen atom of Pro238 no longer exists, and this carbonyl bond is itself reoriented by almost 90° away from the center of the active crevice, as observed in several inhibitor adducts with hydrophobic groups in the S_1' cavity.[27] Overall, the active-site crevice is more open with respect to the previous snapshot (Figure 1c). The observation of a similar adduct with MMP-8[*] (see the Supporting Information) and in a self-interacting form of MMP-12[28] confirms the generality of our findings.

Based on the details of the structures shown in Figure 1a–d and in the corresponding movie, a consistent sequence of events emerges:

1) The uninhibited enzyme has a coordinated water molecule whose position is determined by a strong hydrogen bond with Glu219. The position of this water molecule suggests that the Zn-H_2O-Glu219 moiety has lost at least one proton, as has already been proposed for CPA and TLN. During catalysis the water hydrogen atom in the hydrogen bond to the carboxylate group of Glu219 may actually move to Glu219 as a proton; the coordinated water molecule is thus transformed into a hydroxide ion and its nucleophilicity increased.[29] The active-site crevice is in a "closed" form; that is, it is somewhat narrower than when observed in the presence of the hydrolysis products or inhibitors.
2) The incoming substrate binds zinc with the carbonyl group of *Gly206* (replacing the additional labile water molecules when present), and establishes a number of stabilizing interactions with the protein through its C-terminal section (*IleAlaGly*). It is thought that the N-terminal section may not be involved in relevant interactions with the protein on the basis of the bent conformation observed for the Gln side chain in the product (Figure 1c). The strongly coordinated water molecule (hydroxide ion) performs a nucleophilic attack on the *Gly206* carbonyl group to give the *gem*-diol intermediate modeled in Figure 1b.
3) After the peptide bond has been broken, both peptide fragments remain bound to the protein initially. However, the *ProGlnGly* fragment is only held in place by monodentate coordination of the *Gly206* carboxylate group to the zinc center, whereas the *IleAlaGly* fragment is strongly bound to the protein (Figure 1c). An incoming water molecule loosely binds zinc and separates the bound carboxylate group from the Glu219 residue. The formation of this five-coordinate intermediate facilitates the detachment of the *ProGlnGly* fragment through an associative ligand-exchange mechanism.
4) The remaining protein-bound fragment *IleAlaGly* (Figure 1d) undergoes a relatively small but significant rearrangement in the active-site cavity. The driving force for this rearrangement could rest in the possibly strained pose of this fragment in the two-peptide adduct as a result of several interactions that still exist between the two fragments; these interactions are mediated by the zinc-coordinated water molecule and the Glu219 residue. After the release of the *ProGlnGly* fragment, the increased repulsion between the positively charged zinc ion and the incipient NH_3^+ moiety cause the Ile side chain to enter deeper into the S_1' cavity and optimize its hydrophobic interactions with the cavity. The change in the position of the Ile side chain occurs at the expense of the interaction between the nitrogen atom of *Ala208* and the carbonyl oxygen atom of Pro238; this interaction is lost as a result of the opening up of the cavity. The release of the *IleAlaGly* fragment would from this point follow the same pathway as that of any S_1'-directed inhibitor.

This model may also provide some clues on the pH dependence of the catalytic activity of MMPs. The k_{cat}/K_M profile is bell-shaped, as it is for CPA and TLN, with pK_a values of 4.3 and 9.6. In the absence of any other ionizable groups in the active-site cavity, the simplest interpretation would be that the two pK_a values apply to the successive deprotonations of the coordinated water–Glu219 system. However, mutations in which Glu219 is replaced by a nonionizable residue decrease but do not abolish activity, and, more importantly, the bell-shaped activity profile is maintained.[23] For example, the Glu219Ala mutation in MMP-7 reduces the activity to 0.1%, shifts the first pK_a value to 5.4 (an increase of 1.1 pH units), and does not change the second pK_a value appreciably.[23] In this mutant, the only ionizable residue is the coordinated water molecule, which itself can not account for two pK_a values. One possibility is that one of the

[*] In the MMP-8 adduct the orientation of the *IleAlaGly* fragment is more similar to that observed in the two-peptide adduct of MMP-12, and two water molecules are coordinated to zinc; these water molecules occupy similar positions to those of the two *gem*-diol oxygen atoms in the modeled structure in Figure 1b.

two pK_a values is product-linked. Our mechanism might suggest that the high pK_a value of 9.6 corresponds to the deprotonation of the *IleAlaGly* product. A neutral NH_2 amino terminus could stabilize the *IleAlaGly* adduct, thereby preventing the clearance of the active site for the next enzymatic reaction.

Experimental Section

The catalytic domains of MMP-12 (Gly106–Gly263, F171D mutant) and MMP-8 (Asn85–Gly242) were expressed and purified as reported previously.[30] Crystal growing, data collection, and structure refinement are described in the Supporting Information. The following structures have been deposited to the PDB: Uninhibited MMP-12; the adduct of MMP-12 with two peptides; the adduct of MMP-12 with one peptide; uninhibited MMP-8; the adduct of MMP-8 with one peptide. The PDB file of the model of the MMP-12–*gem*-diol adduct is available as Supporting Information.

Received: July 31, 2006
Published online: November 10, 2006

Keywords: enzyme catalysis · metalloproteins · peptides · reaction mechanisms · X-ray diffraction

[1] S. D. Shapiro, *Curr. Opin. Cell Biol.* **1998**, *10*, 602–608.
[2] A. Boire, L. Covic, A. Agarwal, S. Jacques, S. Sherifi, A. Kuliopulos, *Cell* **2005**, *120*, 303–313.
[3] P. Koolwijk, N. Sidenius, E. Peters, C. F. Sier, R. Hanemaaijer, F. Blasi, V. W. van Hinsbergh, *Blood* **2001**, *97*, 3123–3131.
[4] M. Whittaker, C. D. Floyd, P. Brown, A. J. Gearing, *Chem. Rev.* **1999**, *99*, 2735–2776.
[5] J. W. Skiles, N. C. Gonnella, A. Y. Jeng, *Curr. Med. Chem.* **2004**, *11*, 2911–2977.
[6] N. Borkakoti, *Prog. Biophys. Mol. Biol.* **1998**, *70*, 73–94.
[7] K. S. Larsen, D. S. Auld, *Biochemistry* **1989**, *28*, 9620–9625.
[8] W. L. Mock, J.-T. Tsay, *J. Biol. Chem.* **1988**, *263*, 8635–8641.
[9] D. W. Christianson, J. D. Lipscomb, *Acc. Chem. Res.* **1989**, *22*, 62–69.
[10] D. W. Christianson, W. N. Lipscomb, *J. Am. Chem. Soc.* **1987**, *109*, 5536–5538.
[11] B. L. Vallee, D. S. Auld, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 220–224.
[12] D. G. Hangauer, A. F. Monzingo, B. W. Matthews, *Biochemistry* **1984**, *23*, 5730–5741.
[13] B. W. Matthews, *Acc. Chem. Res.* **1988**, *21*, 333–340.
[14] D. R. Holland, A. C. Hausrath, D. Juers, B. W. Matthews, *Protein Sci.* **1995**, *4*, 1955–1965.
[15] D. H. Juers, J. Kim, B. W. Matthews, S. M. Sieburth, *Biochemistry* **2005**, *44*, 16524–16528.
[16] B. Lovejoy, A. M. Hassell, M. A. Luther, D. Weigl, S. R. Jordan, *Biochemistry* **1994**, *33*, 8207–8217.
[17] L. M. Coussens, B. Fingleton, L. M. Matrisian, *Science* **2002**, *295*, 2387–2392.
[18] N. Borkakoti, *Biochem. Soc. Trans.* **2004**, *32*, 17–20.
[19] V. Lukacova, Y. F. Zhang, M. Mackov, P. Baricic, S. Raha, J. A. Calvo, S. Balaz, *J. Biol. Chem.* **2004**, *279*, 14194–14200.
[20] C. Andreini, L. Banci, I. Bertini, C. Luchinat, A. Rosato, *J. Proteome Res.* **2004**, *3*, 21–31.
[21] J. F. Fisher, S. Mobashery, *Cancer Metastasis Rev.* **2006**, *25*, 115–136.
[22] C. M. Overall, O. Kleifeld, *Br. J. Cancer* **2006**, *94*, 941–946.
[23] J. Cha, D. S. Auld, *Biochemistry* **1997**, *36*, 16019–16024.
[24] A. L. Gall, M. Ruff, R. Kannan, P. Cuniasse, A. Yiotakis, V. Dive, M. C. Rio, P. Basset, D. Moras, *J. Mol. Biol.* **2001**, *307* 577–586.
[25] F. Grams, V. Dive, A. Yiotakis, I. Yiallouros, S. Vassiliou, R. Zwilling, W. Bode, W. Stocker, *Nat. Struct. Biol.* **1996**, *3*, 671–675.
[26] W. N. Lipscomb, N. Strater, *Chem. Rev.* **1996**, *96*, 2375–2433.
[27] H. Brandstetter, F. Grams, D. Glitz, A. Lang, R. Huber, W. Bode, H. W. Krell, R. A. Engh, *J. Biol. Chem.* **2001**, *276*, 17405–17412.
[28] I. Bertini, V. Calderone, M. Fragai, C. Luchinat, S. Mangani, B. Terni, *Angew. Chem.* **2003**, *115*, 2777–2780; *Angew. Chem. Int. Ed.* **2003**, *42*, 2673–2676.
[29] I. Bertini, C. Luchinat, M. Rosi, A. Sgamellotti, F. Tarantelli, *Inorg. Chem.* **1990**, *29*, 1460–1463.
[30] I. Bertini, V. Calderone, M. Cosenza, M. Fragai, Y.-M. Lee, C. Luchinat, S. Mangani, B. Terni, P. Turano, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5334–5339.

Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

69451 Weinheim, Germany

Snapshots of the Reaction Mechanism of Matrix Metalloproteinases

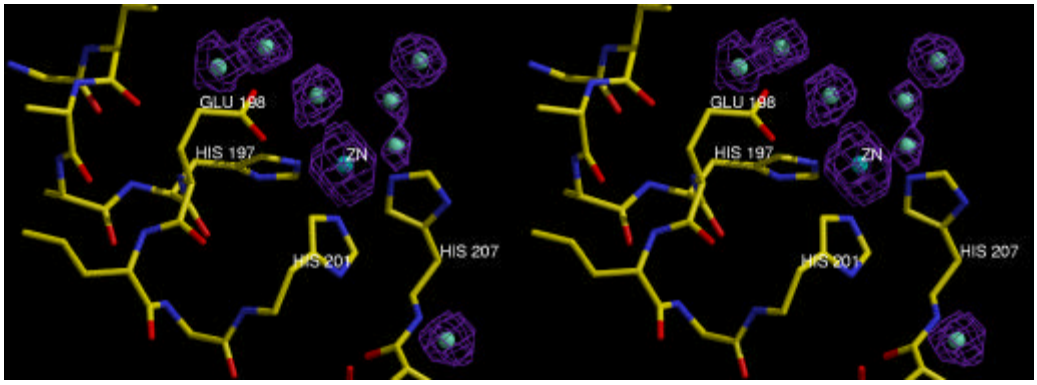
Ivano Bertini,[*,1,2] Vito Calderone,[1] Marco Fragai, [1,3] Claudio Luchinat,[1,3] Massimiliano Maletta,[1,2] Kwon Joo Yeo[1]

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

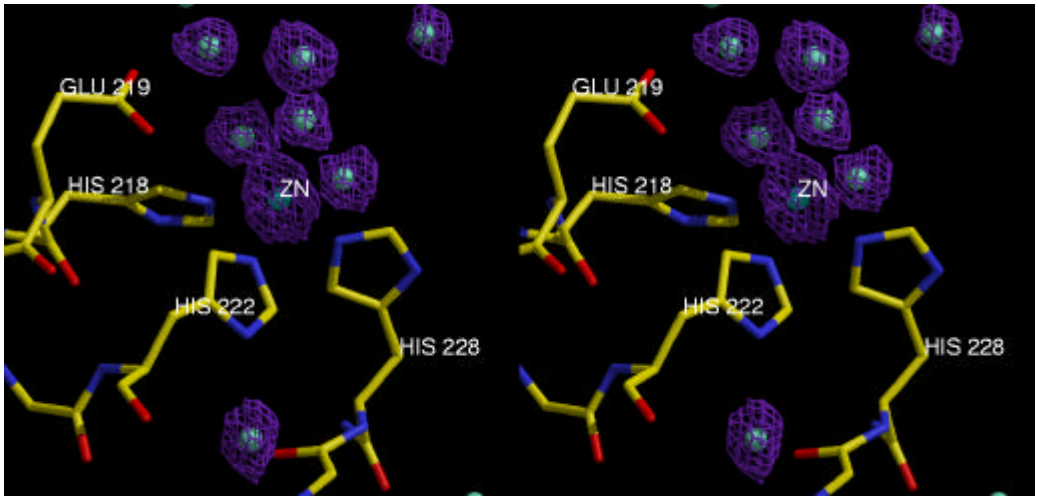
[2] Department of Chemistry – University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Italy

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

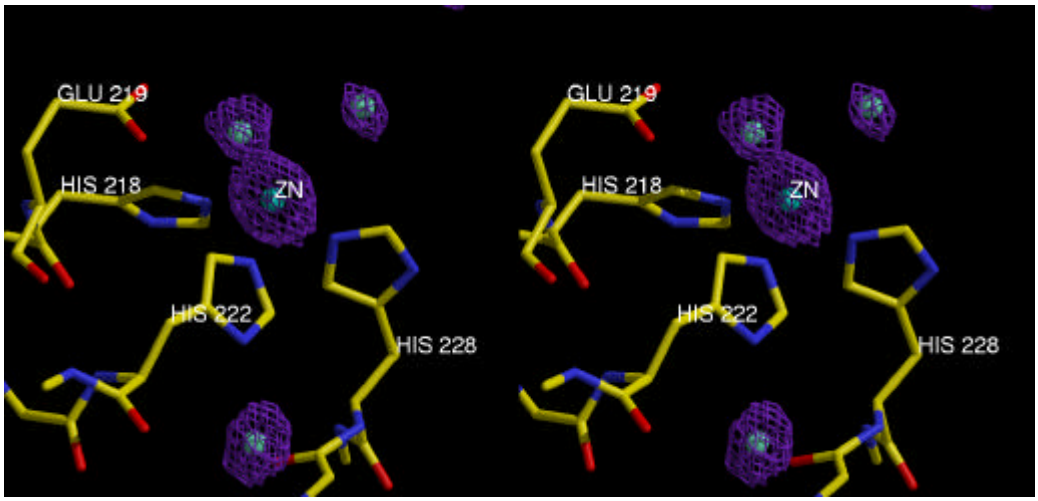
[*]**Prof. Ivano Bertini**
Magnetic Resonance Center (CERM)
University of Florence
Via L. Sacconi 6
50019 Sesto Fiorentino, Italy
e-mail: ivanobertini@cerm.unifi.it
Tel.: +390554574272
Fax: +390554574271



A



B



C

Figure S1. 2Fo-Fc electron density map contoured at 1σ level showing zinc and its coordinated water molecules in **A**) Active MMP-8, **B**) Active MMP-12 and **C**) One peptide MMP-12 adduct.

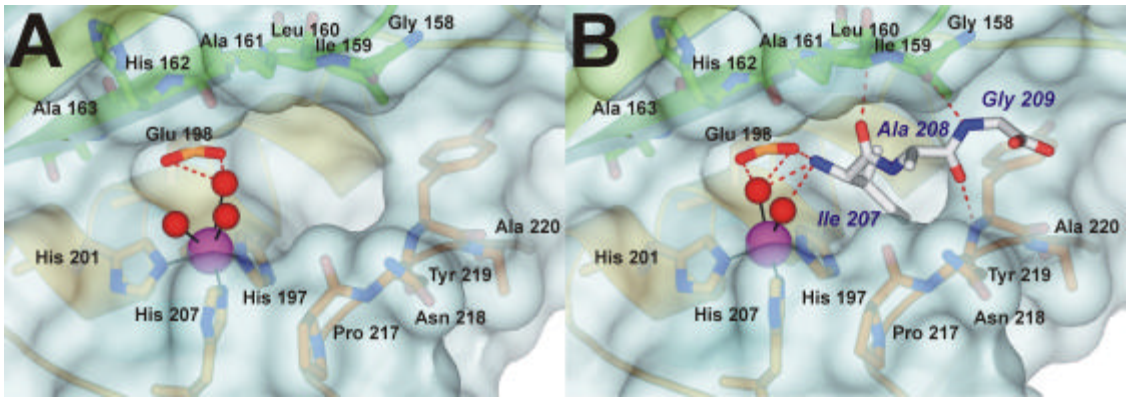
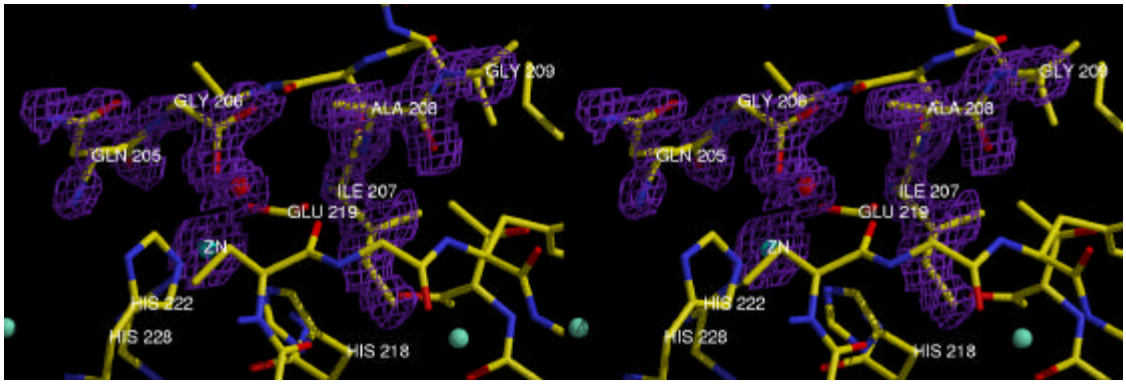
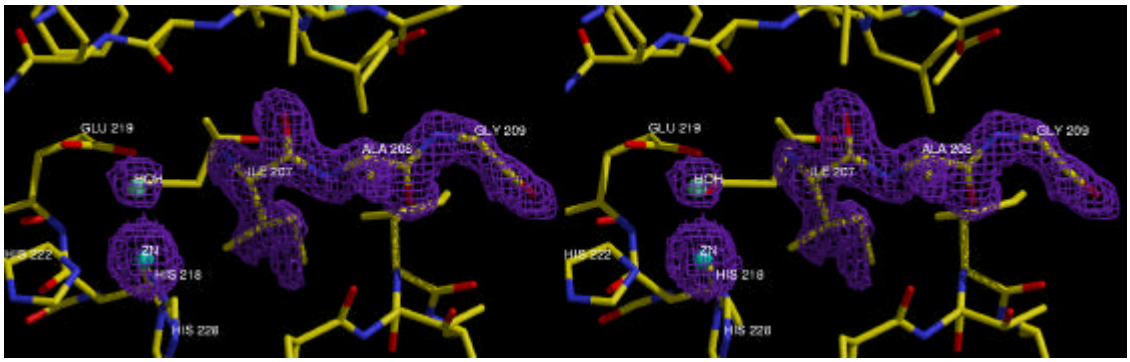


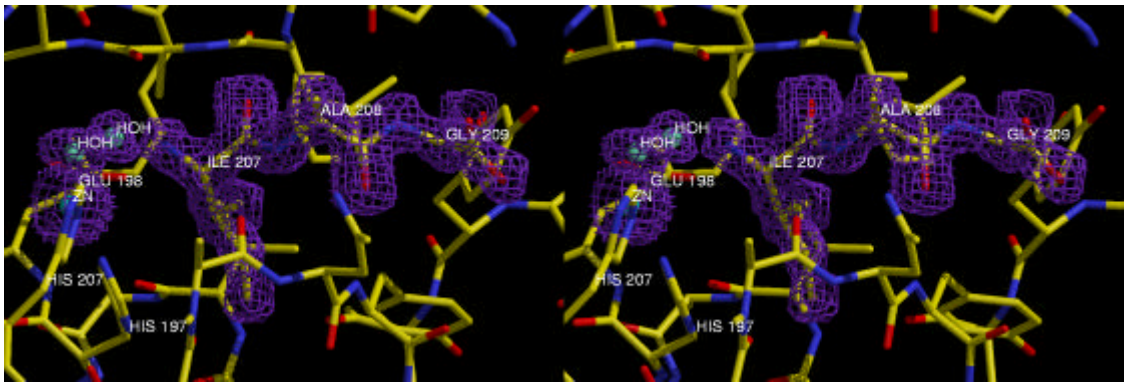
Figure S2. Active, uninhibited form of MMP-8 (A) and its *IleAlaGly* adduct (B).



A

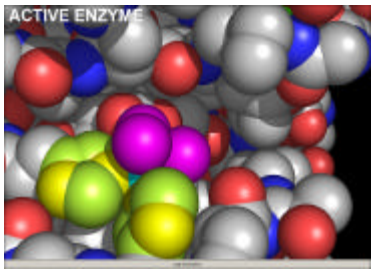


B



C

Figure S3. 2Fo-Fc electron density map contoured at 1σ level showing zinc and its coordinated water molecules and the bound peptide(s) in **A**) Two-peptide MMP-12 adduct, **B**) One peptide MMP-12 adduct and **C**) One peptide MMP-8 adduct.



(set the player *repeat* option on)

Movie S1. A four-frame movie of 1) active MMP, 2) its modelled gem-diol intermediate with the central part of the *ProGlnGlyIleAlaGly* peptide; 3) the two-peptide intermediate; 4) the one-peptide intermediate. The movie is based on the MMP-12 structures, where the gem-diol intermediate[29] in frame (B) is modelled after the uninhibited (Figure 1A) and two-peptide (Figure 1C) forms. Atom colors are CPK except zinc (cyan), its histidine ligands (yellow = N, limon = C), and coordinated water molecules (magenta). Hydrogens are not shown.

Crystallization, Data Collection and Structure Solution

Crystals of human MMP-12 were obtained as previously reported.[1] Crystals of human MMP-8 grew using the same technique at 20° C from a solution containing 0.1 M Tris-HCl, 20% PEG-3350, 200 mM AHA, 0.2 M $MgCl_2$ at pH 8.0. The final protein concentration was 0.4 mM. The crystallization buffer contained 200 mM of the weak inhibitor acetohydroxamic acid (AHA). To obtain the active uninhibited enzymes, MMP crystals were then extensively dialyzed against the same crystallization buffers lacking AHA. The *ProGlnGlyIleAlaGly* peptide (INBIOS s.r.l., Naples) was soaked into the crystals for 1-3 days in order to obtain the two- or the one-peptide adducts.

The peptide was added, in powder form, directly into the drop using a needle and was left incubating for 1-3 days. MMP-12 two-peptide complex (A) was measured in-house, using a PX-Ultra copper sealed tube source (Oxford Diffraction) equipped with an Onyx CCD detector, whereas the single-peptide complex (B) was measured using synchrotron radiation at ID-29 beamline (ESRF, Grenoble, France). Active MMP-12 (C) and the MMP-8 one-peptide complex (D) were measured at beamline BW7B (DESY, Hamburg, Germany), whereas active MMP-8 (E) was measured at beamline ID23-1 (ESRF, Grenoble, France). All datasets were collected at 100 K and the crystals used for data collection were cryo-cooled without any cryo-protectant treatment.

A diffracted to 1.9 Å resolution, B diffracted to 1.2 Å and C to 1.3 Å; they all belong to spacegroup C2 with one molecule in the asymmetric unit, a solvent content of about 50% and a mosaicity of 0.7°-0.8°. D diffracted to 1.5 Å resolution in space group $P2_1$ with two molecules in the asymmetric unit whereas E diffracted to 1.7 Å resolution in space group P1 with two molecules in the asymmetric unit. Solvent content and mosaicity values for D and E are roughly 50% and 0.8°-0.9° respectively.

The data were processed in all cases using the program MOSFLM[2] and scaled using the program SCALA[3] with the TAILS and SECONDARY corrections on (the latter restrained with a TIE SURFACE command) to achieve an empirical absorption correction. Table 1 shows the data collection and processing statistics for all datasets. The structures were solved using the molecular replacement technique; the model used for all MMP-12 datasets was 1Y93 whereas the one used for MMP-8 datasets was 1I73; in all

cases inhibitors, water molecules and ions were omitted from the models. The correct orientation and translation of the molecule within the crystallographic unit cell was determined with standard Patterson search techniques[4,5] (as implemented in the program MOLREP.[6;7] The isotropic refinement was carried out using REFMAC5[8] on A and E datasets but metal ion B-factors were refined taking anisotropy into account; conversely, datasets B, C and D were refined taking anisotropy into account for all atoms. REFMAC5 default weights for the crystallographic term and the geometrical term have been used in all cases.

In between the refinement cycles the models were subjected to manual rebuilding by using XtalView.[9] The same program was used to model ligands. Water molecules have been added by using the standard procedures within the ARP/WARP suite.[10] The stereochemical quality of the refined models was assessed using the program Procheck.[11] The Ramachandran plot is in all cases of very good quality.

Table 1 reports the data collection and refinement statistics for all datasets.

Modelling of the gem-diol

The gem-diol adduct of MMP-12 was modeled using the experimental structures of the active enzyme (Figure 1A) and of the two-peptide adduct (Figure 1C) as the starting point and the gem-diol coordination geometry experimentally observed for a transition state analog.[29]

The model was refined using the *local search* option of Autodock.[12] The standard, validated Autodock zinc parameters for MMP were used.[13;14] A final minimization of both models using Amber[15] converged to the single structure of Figure 1B.

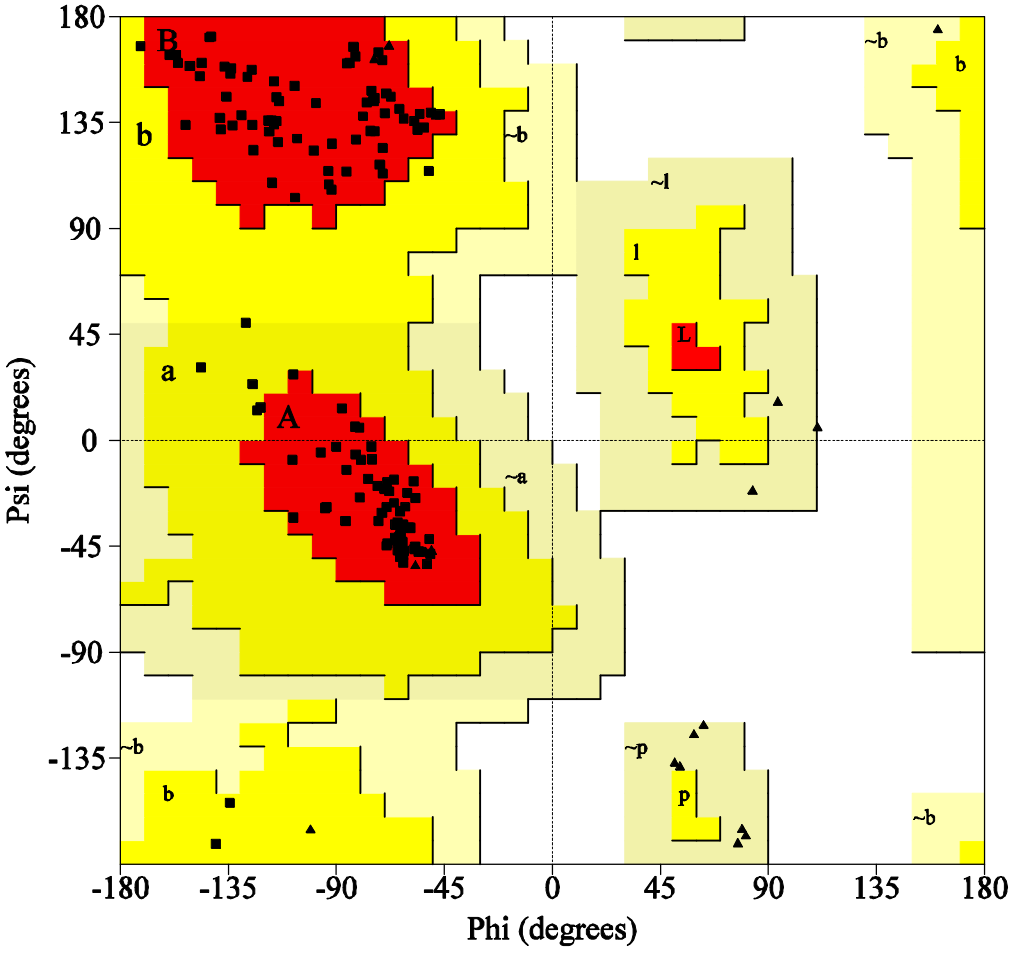
	His218	His222	His228	HOH1 (Glu219)	HOH2	HOH3
Active MMP-12	2.13	2.12	2.09	2.40	2.38	2.10
Two-peptide MMP-12	2.11	2.00	2.02	2.83	2.37	-
One-peptide MMP-12	2.04	2.08	2.04	2.28	-	-
Active MMP-8 (two molecules)	2.06 2.09	2.22 2.06	2.14 2.20	2.84 2.69	2.70 -	2.79 -
One-peptide MMP-8 (two molecules)	2.07 2.06	2.10 2.12	2.04 2.00	2.27 2.19	2.35 2.48	- -

Table S1. Distances between Zn and the three coordinated histidines, between the Glu-activated water molecules and zinc and between Zn and the other water molecules coordinated to it.

Table S2.

Table 1.	**DATA COLLECTION AND REFINEMENT STATISTICS**				
	Two-peptide-MMP12 complex (A)	Active MMP-12 (C)	One-peptide-MMP12 complex (B)	Active MMP-8 (E)	One-peptide-MMP8 complex (D)
Spacegroup	C2	C2	C2	P1	$P2_1$
Cell dimensions (Å, °)	a= 51.54 b= 60.37 c= 54.45 β= 115.41	a= 51.54 b= 60.75 c= 54.26 β= 115.64	a= 51.89 b= 60.36 c= 54.52 β= 115.73	a= 33.29 b= 47.11 c= 61.32 α= 77.73 β= 80.03 γ= 77.01	a= 33.21 b= 68.53 c= 78.28 β= 98.10
Resolution (Å)	30.2 – 1.9	25.8 – 1.2	49.0 – 1.1	39.5 – 1.7	38.7 – 1.5
Unique reflections	11726 (1505)	41969 (5315)	50850 (7153)	30329 (4603)	55607 (8108)
Overall completeness (%)	98.1 (86.9)	97.2 (84.8)	94.6 (91.1)	90.3 (88.3)	99.9 (99.9)
R_{sym} (%)	13.7 (32.7)	7.9 (12.6)	5.9 (15.6)	10.3 (24.5)	8.2 (31.4)
Multiplicity	5.6 (3.0)	5.8 (5.2)	6.8 (6.8)	1.5 (1.5)	6.7 (6.0)
I/(σI)	4.7 (2.3)	5.1 (3.8)	4.2 (4.1)	4.7 (2.6)	6.6 (2.2)
Wilson plot B-factor ($Å^2$)	7.69	7.17	11.08	16.14	10.21
R_{cryst} / R_{free} (%)	20.6 / 28.7	19.6 / 21.6	19.7 / 22.2	22.5 / 29.3	16.3 / 19.2
Protein atoms	1238	1238	1238	2480 (two molecules)	2480 (two molecules)
Ions	5	5	5	8	8
Ligand atoms	31	0	17	0	17
Water molecules	119	238	206	303	591
RMSD bond lengths (Å)	0.021	0.007	0.007	0.020	0.008
RMSD bond angles (°)	1.8	1.0	1.1	1.6	1.1
Mean B-factor ($Å^2$)	9.80	11.00	14.16	17.17	12.39

PROCHECK



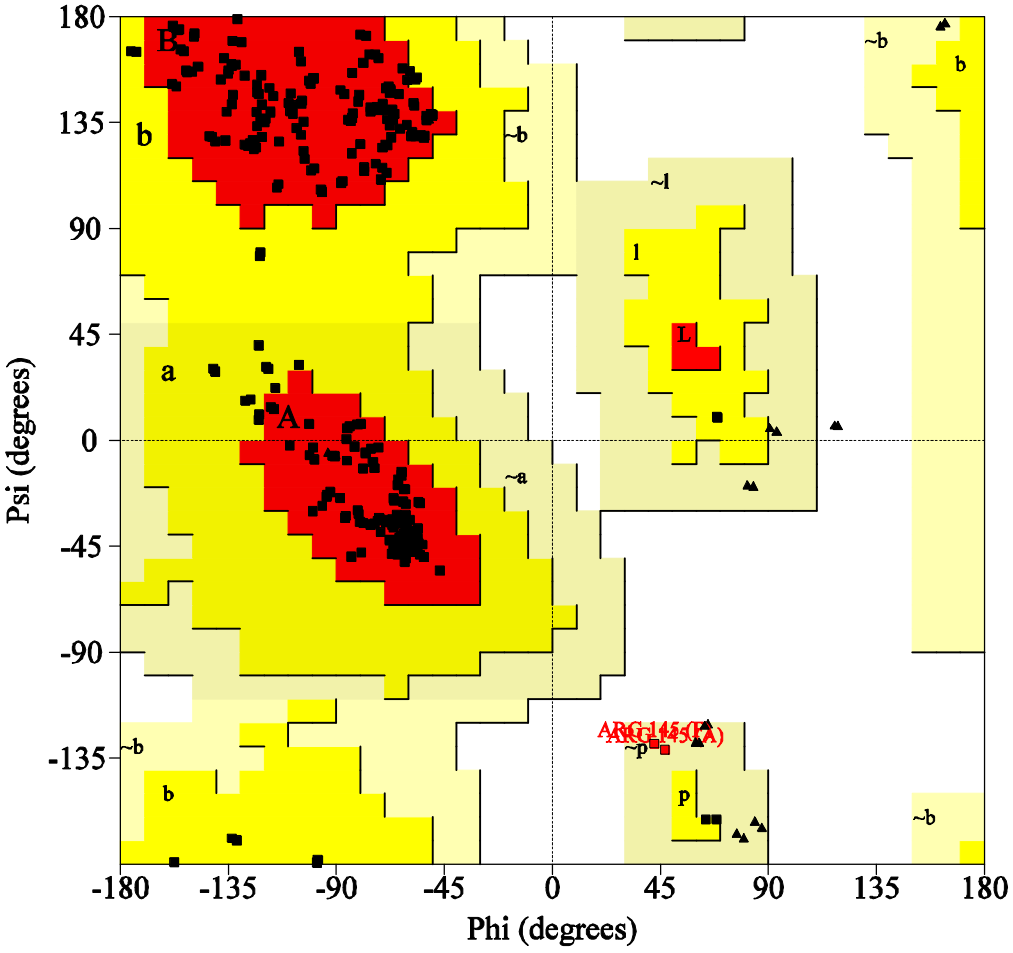
Plot statistics

Residues in most favoured regions [A,B,L]	124	93.2%
Residues in additional allowed regions [a,b,l,p]	9	6.8%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	133	100.0%
Number of end-residues (excl. Gly and Pro)	240	
Number of glycine residues (shown as triangles)	19	
Number of proline residues	6	
Total number of residues	398	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

A

PROCHECK



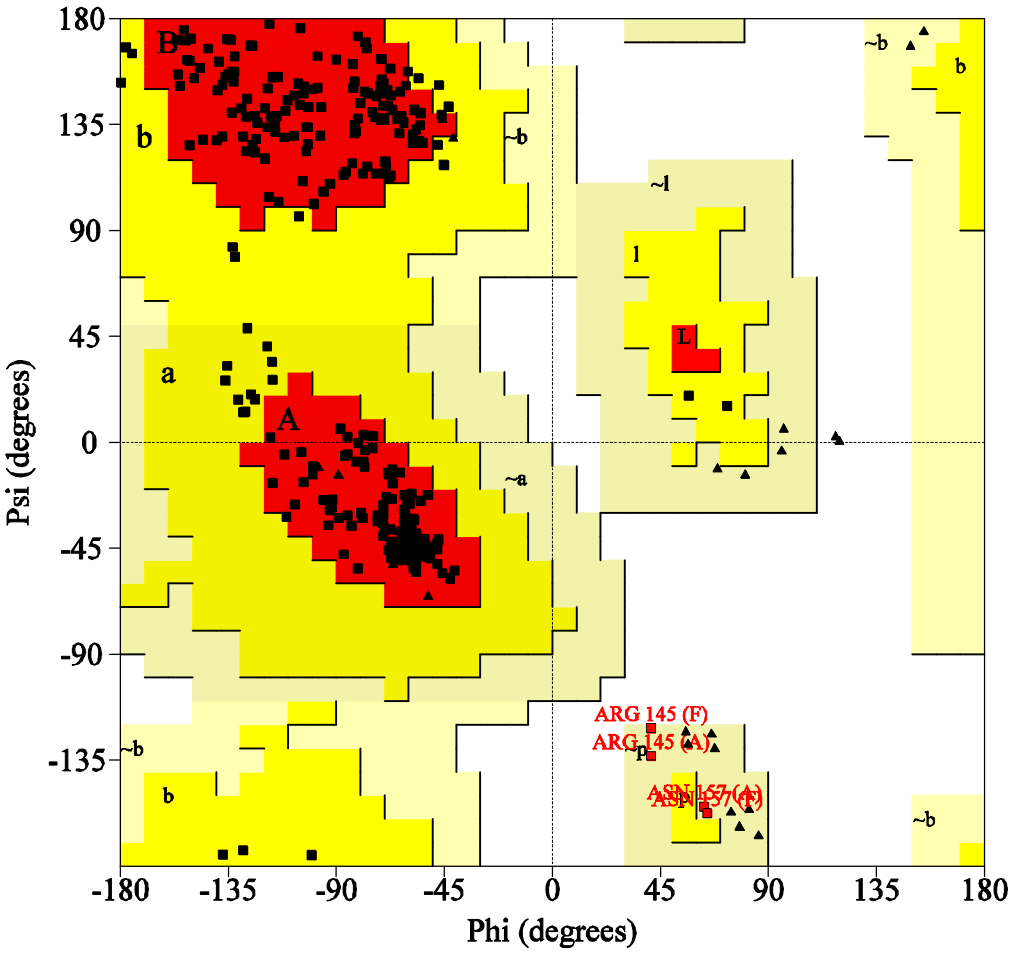
Plot statistics

Residues in most favoured regions [A,B,L]	245	90.1%
Residues in additional allowed regions [a,b,l,p]	25	9.2%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.7%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	272	100.0%
Number of end-residues (excl. Gly and Pro)	575	
Number of glycine residues (shown as triangles)	28	
Number of proline residues	18	
Total number of residues	893	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

B

PROCHECK



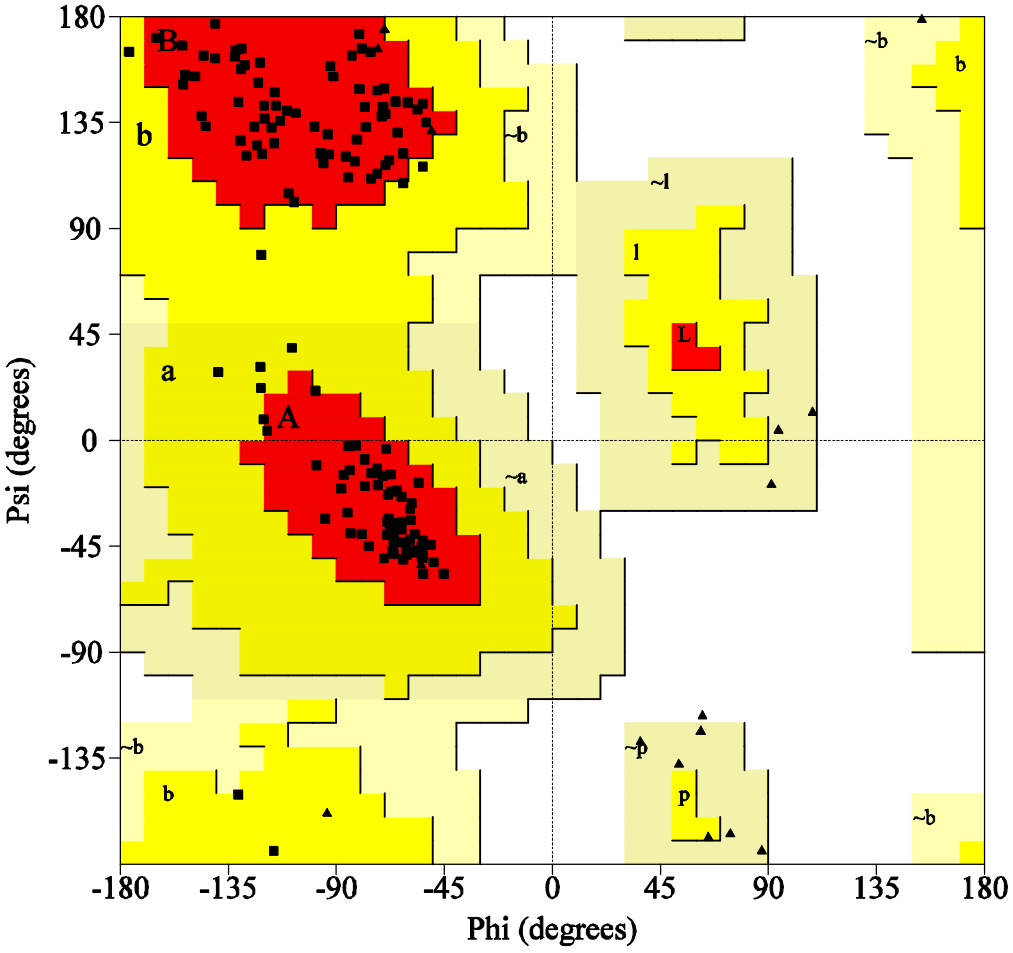
Plot statistics

Residues in most favoured regions [A,B,L]	241	89.3%
Residues in additional allowed regions [a,b,l,p]	25	9.3%
Residues in generously allowed regions [~a,~b,~l,~p]	4	1.5%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	270	100.0%
Number of end-residues (excl. Gly and Pro)	304	
Number of glycine residues (shown as triangles)	26	
Number of proline residues	18	
Total number of residues	618	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

C

PROCHECK



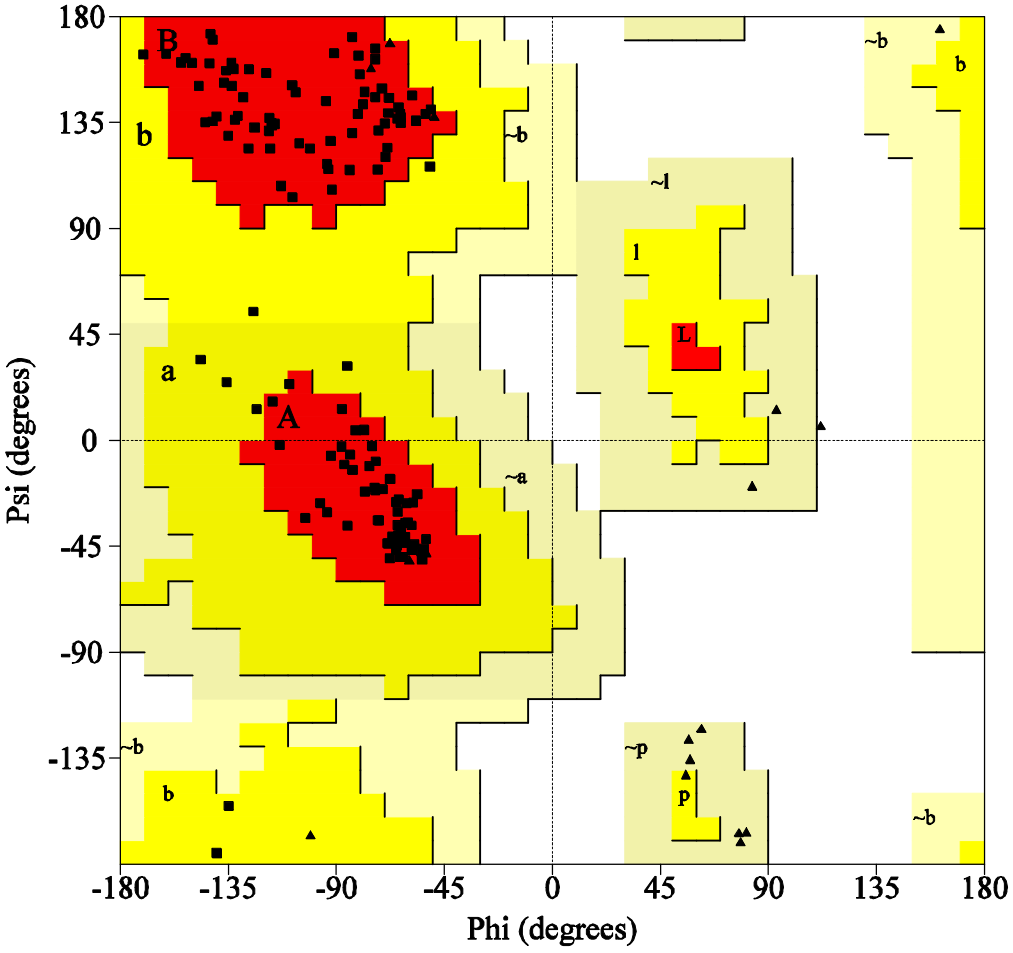
Plot statistics

Residues in most favoured regions [A,B,L]	122	91.0%
Residues in additional allowed regions [a,b,l,p]	12	9.0%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	134	100.0%
Number of end-residues (excl. Gly and Pro)	115	
Number of glycine residues (shown as triangles)	21	
Number of proline residues	6	
Total number of residues	276	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

D

PROCHECK



Plot statistics

Residues in most favoured regions [A,B,L]	126	94.0%
Residues in additional allowed regions [a,b,l,p]	8	6.0%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
Number of non-glycine and non-proline residues	134	100.0%
Number of end-residues (excl. Gly and Pro)	205	
Number of glycine residues (shown as triangles)	20	
Number of proline residues	6	
Total number of residues	365	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

E

Figure S4. Ramachandran plots for A) Active MMP-12, B) One peptide MMP-8 adduct, C) Active MMP-8, D) Two peptide MMP-12 adduct and E) One peptide MMP-12 adduct.

References

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

[2] A. G. W. Leslie, in *Molecular data processing*Eds.: D. Moras, A. D. Podjarny, J.-C. Thierry, Oxford University Press, Oxford **1991**, 50-61.

[3] P. R. Evans, "Data Reduction", Proceedings of CCP4 Study Weekend, *"Data Reduction", Proceedings of CCP4 Study Weekend*, pp. 114-122 .

[4] M. G. Rossmann, D. M. Blow, *Acta Cryst.* **1962,** *D15*, 24-31.

[5] R. A. Crowther, Ed.: M. G. Rossmann), Gordon & Breach, New York **1972**.

[6] A. Vagin, A. Teplyakov, *J.Appl.Crystallogr.* **1997,** *30*, 1022-1025.

[7] A. Vagin, A. Teplyakov, *Acta Crystallogr D Biol Crystallogr* **2000,** *56*, 1622-1624.

[8] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Cryst.* **1997,** *D53*, 240-255.

[9] D. E. McRee, *J.Mol.Graphics* **1992,** *10*, 44-47.

[10] V. S. Lamzin, *Acta Crystallogr D Biol Crystallogr* **1993,** *49*, 129-147.

[11] R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, *J.Appl.Crystallogr.* **1993,** *26*, 283-291.

[12] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J Comp Chem* **1998,** *19*, 1639-1662.

[13] X. Hu, W. H. Shelver, *J.Mol.Graph.Model.* **2003,** *22*, 115-126.

[14] I. Bertini, M. Fragai, A. Giachetti, C. Luchinat, M. Maletta, G. Parigi, K. J. Yeo, *J.Med.Chem.* **2005,** *48*, 7544-7559.

[15] Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., et. al. AMBER 8, 2004. San Francisco, CA, University of California.

4.5 Exploring the Subtleties of Drug-Receptor Interactions: the Case of Matrix Metalloproteinases [Accepted in *J. Am. Chem. Soc*.]

Exploring the Subtleties of Drug-Receptor Interactions: the Case of Matrix Metalloproteinases

Bertini Ivano,[*,1,2] Calderone Vito,[1] Fragai Marco,[1,3] Giachetti Andrea,[1,3,4] Loconte Mauro,[4] Luchinat Claudio,[1,3] Maletta Massimiliano,[1,2] Nativi Cristina,[5] Kwon Joo Yeo[1]

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

[2] Department of Chemistry – University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Italy

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

[4] ProtEra S.r.l. , Via delle Idee, 22, 50019 Sesto Fiorentino (Italy)

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

[*]**Prof. Ivano Bertini**
Magnetic Resonance Center (CERM)
University of Florence
Via L. Sacconi 6
50019 Sesto Fiorentino, Italy
e-mail: ivanobertini@cerm.unifi.it
Tel.: +390554574272
Fax: +390554574271

Abstract

By solving high resolution crystal structures of a large number (14 in this case) of adducts of matrix metalloproteinase 12 (MMP12) with strong, nanomolar, inhibitors all derived from a single ligand scaffold, it is shown that the energetics of the ligand-protein interactions can be accounted for directly from the structures to a level of detail that allows us to rationalize for the differential binding affinity between pairs of closely related ligands. In each case, variations in binding affinities can be traced back to slight improvements or worsening of specific interactions with the protein of one or more ligand atoms. Isothermal calorimetry measurements show that the binding of this class of MMP inhibitors is largely enthalpy driven, but a favorable entropic contribution is always present. The binding enthalpy of acetohydroxamic acid (AHA), the prototype zinc-binding group in MMP drug discovery, has been also accurately measured. In principle, this research permits the planning of either improved inhibitors, or inhibitors with improved selectivity for one or another MMP. The present analysis is applicable to any drug target for which structural information on adducts with a series of homologous ligands can be obtained, while structural information obtained from *in silico* docking is probably not accurate enough for this type of study.

Introduction

Drug discovery projects aim at finding small molecules that bind a selected protein target with high binding affinity. High selectivity is also a desirable property especially when the selected target belongs to a protein family. In the case of matrix metalloproteinases (MMP), a family of closely related zinc-containing extracellular proteases, numerous high affinity inhibitors are available.[1] Well-known examples are batimastat and marimastat belonging to the class of succinate peptidomimetic inhibitors,[2,3] the class of succinate macrocyclic inhibitors,[4] the class of sulfone hydroxamate,[5] the class of the reverse hydroxamic acid inhibitors[6] and the class of sulfonamidic inhibitors such as NNGH.[7] Their dissociation constants are in the low nanomolar range. However, the selectivity versus one or another MMP is often modest and this is at the origin of several side-effects discovered during the clinical trials.[8] Several approaches have been proposed in order to design selective ligands for protein targets.[9-12] Three-dimensional structures of enzyme-inhibitor adducts for several ligands, belonging to the various classes, with several different MMPs are available.[3,5,13-22] 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 (Figure 1A). In such a situation, to attempt planning ligands with increased selectivity requires an understanding of the energetics of the inhibitor interactions with each of these active site regions to an unprecedented level of detail. The existing structural data for various MMP-inhibitor adducts are many but scattered, and structure-affinity relationships at the level of a meaningful dissection of the various contributions can be hardly attempted.[23] Furthermore, the catalytic domains of MMPs undergo non-negligible conformational equilibria in solution,[17,24,25] which also involve the active site pocket. A flexible active site is a serious drawback for the use of *in silico* docking programs to design inhibitors,[26] and even more to refine them to increase selectivity.[27]

With this in mind, we felt it would be important to ascertain whether, in case a large number of structures of MMP adducts with a homologous series of ligands were

available, structure-affinity relationships can be established; and, if so, whether these relationships are accurate enough to enable meaningful protein-specific ligand-refinement strategies. Thus, we collected structural data for a large number of adducts of a homologous series of ligands with one particular MMP. We selected MMP12, a validated target for emphysema[28] and multiple sclerosis,[29] because a crystalline form of its catalytic domain is available[17] that allows easy soaking of inhibitors. High-resolution crystal structures of as many as sixteen different adducts of the protein with structurally related inhibitors were obtained. In parallel, the free energies of binding (ΔG^0) through enzyme activity inhibition experiments were measured for the same ligands. It is shown that, indeed, a careful analysis of the structural data, performed with the aid of structure-based computed interatomic interaction energies, can provide a detailed description of how different substituents on the same scaffold can interplay to improve or worsen affinity and selectivity. It is proposed that this strategy is of general validity for any drug target, provided structural information on a number of homologous ligand-protein adducts can be obtained.

Materials and Methods

Synthesis, characterization and inhibitory assays of compounds **1-11** and **13-14** will be described elsewhere. Compound 12 has been already described.[7]

Expression and purification of human MMP12 and MMP13 catalytic domain. Cloning, expression and purification of MMP12 catalytic domain have been previously described.[17] The cDNA of proMMP13 (Leu20-Pro268) was cloned into the pET21 vector (Novagen) using *Nde I* and *Xho I* as restriction enzymes. One additional methionine at position 19 was present in the final expression product. The *E. coli* strain BL21 Codon Plus cells, transformed with the above plasmid, were grown in LB medium at 310 K. The protein expression was induced during the exponential growth phase with 0.5 mM of IPTG. The cells were harvested 3 hours after induction. After lysis of the cells the inclusion bodies, containing the proMMP13, were solubilized in 6 M urea; 20 mM Tris (pH 8.0). The protein was

purified on the Hitrap Q column (Pharmacia) with a buffer containing 6 M urea; 20 mM Tris (pH 8.0). The elution was performed using a linear gradient of NaCl up to 0.6 M. The purified protein was then refolded using multi-step dialysis against solutions containing 50 mM Tris (pH 7.2); 10 mM $CaCl_2$; 0.1 mM $ZnCl_2$; 0.3 M NaCl and decreasing concentration of urea. The refolded protein was exchanged, by dialysis, against a buffer with 50 mM Tris (pH 7.2); 5 mM $CaCl_2$; 0.1 mM $ZnCl_2$; 0.3 M NaCl. The protein was concentrated at room temperature using an Amicon up to a concentration of about 30 μM. The active protein is left overnight in these conditions to allow the autoproteolysis of the prodomain. After addition of Acetohydroxamic acid (AHA) to a final concentration of 0.5 M, the catalytic domain of MMP13 (Tyr 104- Pro 268) was purified using size-exclusion chromatography with the buffer 50 mM Tris (pH 7.2); 5 mM $CaCl_2$; 0.1 mM $ZnCl_2$; 0.3 M NaCl, 0.2 M AHA and then concentrated using a Centriprep concentrator at 277 K to a final concentration of 0.3 mM.

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.

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 $LiCl_2$. Crystals of MMP12-AHA were obtained in the presence of $LiCl_2$ as well.[17]

The data were measured in-house, using a PX-Ultra copper sealed tube source (Oxford Diffraction), for **3**, **6**, **7**, **9**, **13** and **14**, at ID29 ESRF (Grenoble, France) for **4** and **8**, at BW7A DESY (Hamburg, Germany) for **1**, **2**, **5**, **10** and **11** complexes and at XRD-1 ELETTRA (Trieste, Italy) for **12**.

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 adducts crystallize in the C2 space group with one molecule in the asymmetric unit while co-crystallized complex of **12** 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[30] and scaled using the program SCALA[31] with the TAILS and SECONDARY corrections on (the latter restrained with a TIE SURFACE command) to achieve an empirical absorption correction. Table 1S (in Supporting Information) shows the data collection and processing statistics for all datasets.

The structure of the adduct with **12** was previously solved[17] 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[32,33] as implemented in the program MOLREP.[34,35] The refinement was carried out using REFMAC5[36] 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.[37] 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[38] 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.[39] The Ramachandran plot for all structures is of very good quality.
The coordinates for all adducts are under deposition at the ProteinDataBank.

Energy calculation. Autodock 3.05 was used to calculate the protein ligand binding-energy.[40] The calculation was performed starting from the x-ray crystal structure using reliable Zn(II) parameters.[41] A box of 70x70x70 point with grid spacing of 0.375 Å was created on the protein side, and centered near the catalytic Zn(II) ion. The ligands were extracted from the x-ray structure and hydrogen atoms were added considering the hydroxamic group deprotonated; Gaister-Marsili partial atomic charges were calculated using BABEL.[42] The ligands were modelled based on the X-ray structures which were already minimized against the experimental diffraction data. Energy contributions for individual atoms were extracted from the Autodock energy output.

Calorimetry. Isothermal Titration Microcalorimetry experiments were performed at 298 K with a VP-ITC microcalorimeter (MicroCal, Inc., Northampton, MA). After an initial injection of 1 μL, aliquots of 9 μL of 200 μM inhibitor were stepwise injected into the sample cell containing a solution 20 μM of MMP12 catalytic domain until complete saturation. All experiments were performed in 20 mM Tris (pH 7.2), 5 mM $CaCl_2$, 0.1 mM $ZnCl_2$, 0.3 M NaCl, AHA 4 mM with 0.1% (v/v) DMSO. Heats of dilution were measured by injecting the ligand into buffer and then subtracted from the binding heats. The thermodynamic parameters and K_A values were calculated fitting data to a single binding site model with ORIGIN 7.0 sofware (Microcal, Inc.).[43]

Results and Discussion

Crystal structures and dissociation constants of ligand-MMP12 adducts. The crystal structures of the complexes of the catalytic domain of MMP12 with 14 closely related inhibitors have been solved. Figure 1B shows the scaffold common to all these inhibitors, and the three positions on which different functionalizations have

been tested. The general formula contains a hydroxamic group targeted to bind the catalytic zinc ion, and an aromatic moiety bearing the R_1 substituent targeted to the hydrophobic S_1' pocket. The two groups are tethered by a sulphonamide moiety interacting with the substrate binding groove and containing substituents R_2 and R_3, potentially able to give rise to further interactions. R_1 ranges from H to F to OCH_3 to C_6H_5. These variations were planned to monitor their relative effects on the strength of the interaction with S_1'. R_2 and R_3 are either H or aliphatic groups with various degrees of hydrophobicity/hydrophilicity.

A close up of the arrangement of the various ligands in the active site of MMP12 is shown in Figures 2-6. All the inhibitors share the same binding mode, as expected, and to a first inspection all relevant interactions (with the S_1' pocket, the metal and the substrate binding groove) are in place. To emphasize the similarities, all structures are reported with the same orientation, obtained by least square fits of all the coordinates of the adducts to one another. The left-hand column of Figures 2-6 shows the structures of each adduct, and the central column reports the structure of the adduct with the closest ligand homologue.

Despite the similarities of the ligand-protein interactions, the dissociation constants, measured as inhibition constants through fluorimetry under non-saturating substrate concentration conditions at 298 K (see materials and methods), span two orders of magnitude. The associated ΔG^0 values are reported in Table 1.[†] It can be seen that there are relevant differences in binding affinities even between ligands that are very close homologues, *i.e.* are next to one another in Figures 2-6.

These results illustrate well the nature of the problem: as large differences in the binding modes can be immediately ruled out by visual inspection, the point is to ascertain whether small differences in the binding mode are at all detectable[‡] and, if

[†]As described in the Materials and methods Section, the protein was used as an adduct with the weak inhibitor acetohydroxamic acid (AHA). Stock protein solutions contained 200 mM AHA. At the final dilution for the fluorimetric assays (1 nM of enzyme), the AHA concentration was 5 μM at most, which is well below the dissociation constant of 6.2 mM estimated through the same fluorimetric assay. Therefore, it may be concluded that AHA does not appreciably interfere with the assays.

[‡] The nominal resolution for the adduct structures obtained in this work range between 1.2 and 1.7 Å, which corresponds to indeterminations in interatomic distances of ±0.10 to ±0.15 Å.

so, whether these small differences are consistently related with the differences in binding affinity.

Careful examination of key interatomic distances and angles does suggest that indeed not all interactions, despite being all in place, may be energetically equivalent. For example, on passing from R_1 = -H (**1**) to R_1 = -F (**2**), -OCH_3 (**3**), -C_6H_5 (**4**), Van der Waals contacts are of course gained in the S_1' pocket, but less optimal H-bond interactions of the SO_2 group are seen in all cases, a slight worsening is apparent in the hydroxamic acid coordination mode in the -OCH_3 derivative, and a less deep penetration of the first phenyl ring in the S_1' pocket is seen in the -C_6H_5 derivative. As another example, by comparing the R_1 = -C_6H_5, R_2 = -CH_2CH_2OH, R_3 = -H derivative (**7**) with the R_1 = -C_6H_5, R_2 = -H, R_3 = -H derivative (**4**) on one side, and the R_1 = -OCH_3, R_2 = -CH_2CH_2OH, R_3 = -H derivative (**6**) with the R_1 = -OCH_3, R_2 = -H, R_3 = -H derivative (**3**) on the other side, we notice that the addition of the R_2 = -CH_2CH_2OH substituent may have opposite effects: it modestly improves the overall binding of the R_1 = -OCH_3 compound, and strongly worsens the binding of the R_1 = -C_6H_5 compound. Inspection of the structures suggests that the worsening of the affinity of the -C_6H_5 derivative may be qualitatively explained by the less optimal binding mode of the hydroxamic moiety, with an average distance of the two oxygens of 0.03 Å longer, respectively, in the -C_6H_5 with respect to the -OCH_3 case. While an increase in binding affinity upon increasing the R_1 substituent in the first example could have been expected, the reversal in behavior of the R_2 = -CH_2CH_2OH substitution between the -OCH_3 and the -C_6H_5 derivatives was unexpected.

Differential interaction energies of homologous ligand pairs. From careful inspection of the structures a qualitative correlation seems to exist between good intermolecular interactions and binding affinity. It is difficult, however, to extend these considerations to all molecules, especially when some interactions seem to have improved and other to have been worsened in the same adduct. In other words, it would be desirable to have a simpler but reliable way to translate these observations in contributions to binding. As the individual interaction strengths can be ultimately

traced down to either hydrophobic or electrostatic effects, we decided to simply evaluate the hydrophobic and electrostatic contributions to the overall intermolecular interaction energy *by each ligand atom* using standard parameter sets such as those employed in popular docking programs. Among them, the parameter set used by Autodock was selected because of its simple *ad hoc* treatment of the ligand metal ion interaction.[44] It should be immediately stressed that we are not using these parameters to estimate overall, absolute binding energies (see however later) but only to attach a more quantitative meaning to structural observations by calculating *differences* in interatomic interaction energies between the protein and each atom of different homologous ligands. By estimating differences, possible biases originating by crude approximations, for instance those used by Autodock to parameterize the strength of the zinc-hydroxamic ligand coordination bonds, should largely cancel. Furthermore, the differences are taken between ligands that are very close homologues, *i.e.* are next to one another in Figures 2-6.

The results are reported in pictorial form on the right-hand columns of Figures 2-6. For each row in these figures, the ligand forming the adduct reported in the left-hand side of the figure is drawn, and its atoms are color- and size-coded to show whether their individual interactions with the protein are stronger (green) or weaker (red), and by how much (sphere volume), with respect to those of the corresponding atoms in the central column adduct. For ligand atoms that do not have a corresponding atom in the homologue compound their interaction strengths with the protein are encoded as such. The right-hand columns of Figures 2-6 thus provides an immediate perception of which atoms may be relevant in determining the differences in affinity when the structures of closely related ligands are compared. We have checked that the results of the above procedure are largely independent of the set of parameters used, suggesting that the estimated differences in interaction energies are reliable. A better way to check their reliability is of course to compare the algebraic sum of all the differences with the difference in experimental ΔG^0 values ($\Delta\Delta G^0$). This comparison is reported in Figure 7. Remarkably, not only are the signs of the

$\Delta\Delta G^0$s predicted correctly in most cases, but their magnitudes are also reproduced quite well. In only two cases (**2-1** and **10-3**) the calculated sign is opposite to the predicted one, but in both cases the differences are rather close to zero. An entropic penalty may be further added to the algebraic sum of the individual energetic terms if, between the two ligands, there is a difference in the number of rotatable bonds that can be possibly immobilized upon binding to the protein. For instance, in Autodock the entropic penalty is assigned a default value of 0.31 kcal/mol per bond. The calculated $\Delta\Delta G^0$ in Figure 7 take this contribution into account when deemed present, *i.e.* when the ligands to be compared differ in number of potentially immobilized rotatable bonds. The bars extending, inwards or outwards of the calculated $\Delta\Delta G^0$ values, show the results obtained by neglecting the entropic penalty. It is apparent that the general trend is maintained whether or not such contribution is taken into account.

In our opinion, the results in Figure 7 constitute a strong validation of the use of energetic considerations to examine the details of the structure-affinity relationships in this series of homologous MMP inhibitors. We turn now to describe such details, to further show that their information content is high enough to make this approach precious in terms of predictive power for the refinement of inhibitors towards a particular target.

R_1 groups. The first three pairs of experimental $\Delta\Delta G^0$ values in Figure 7 show that the substitution of a hydrogen atom in *para* position to the phenyl substituent targeted to the hydrophobic S_1' pocket with –F (**2**), -OCH_3 (**3**) or a -C_6H_5 (**4**) R_1 group is essentially neutral, moderately advantageous, and strongly advantageous, respectively. The corresponding estimated $\Delta\Delta G^0$ values (Figure 7) agree very well with the observations, albeit with a slight tendency to overestimate the improvement. From Figure 2A-C it appears that the atoms involved in improving the binding are mostly the additional atoms in the R_1 group, as expected. The rest of the molecular scaffold is scarcely affected. However, it is interesting to notice that the perturbations on the rest of the scaffold are larger in the -OCH_3 than in the -F derivative, and that

they extend as far as to affect the coordination of the hydroxamic group to the zinc ion. It is also interesting to note that individual interactions can both gain and lose strength, with a prevalence of the latter. This is to be expected as, the bulkier the R_1 group, the fewer the degrees of freedom available to the rest of the molecule to optimize these individual interactions. The worsening of a subset of other interactions in the rest of the molecule is particularly apparent for R_1 = -C_6H_5: its presence does increase substantially the total hydrophobic interactions in the S_1' pocket, but at the same time decreases the hydrophobic interactions of the first phenyl ring (right-hand side of Figure 2C). As noticed already by comparing the left and central panels in Figure 2C, the presence of the second ring apparently prevents the first phenyl ring to enter the S_1' pocket as deeply as in the unsubstituted derivative. Also the other qualitative observations made by comparing the structures in panels A-C in Figure 2 are confirmed: -F introduces the least perturbations, slightly worsening the hydrophobic interactions in the S_1' pocket and the hydrogen bonds of the SO_2 group, while in the -OCH_3 derivative an appreciable worsening is transmitted to the hydroxamic moiety, and in the -C_6H_5 derivative there is a modest but clear decrease of the hydrophobic interactions of the first phenyl ring with the S_1' pocket.

R_2 = -*CH_2CH_2OH*. For three derivatives with R_1 = -F (**5**), -OCH_3 (**6**) and -C_6H_5 (**7**), the effect of substituting the hydrogen atom in position R_2 with -CH_2CH_2OH was investigated. The alcoholic substituent in R_2 slightly improves the binding to the enzyme for the first two derivatives, and sizably worsens the binding for the third. Again, the estimated $\Delta\Delta G_0$ values agree very well with the observations (Figure 7). Panels A-C in Figure 3 provide a rationale for the different behavior. In summary, in all three cases the R_2 substituent provides an appreciable contribution to binding, especially through hydrophobic interactions of the two additional CH_2 groups. For R_1 = -F this is accompanied by a modest worsening of the hydroxamic coordination strength. Conversely, for R_1 = -OCH_3 the hydroxamic coordination is actually slightly improved, and a modest worsening of the hydrophobic contact of the phenyl ring is seen. Thus, in no case a perfect fit is achieved. It is interesting to note that the non

perfect fits for the two different molecules lead to different compromises for the interactions with the three enzyme subsites: For R_1 = -F, the interaction is optimized for the hydrophobic pocket and the substrate binding groove, while the binding to zinc is sacrificed. For R_1 = -OCH_3 the interaction is optimized for the binding to zinc and to the substrate binding groove, while the binding to the hydrophobic pocket is sacrificed.

Unexpectedly, the R_2 = -CH_2CH_2OH substitution has a sizable destabilizing effect on the R_1 = -C_6H_5 scaffold. In this case, the binding to zinc is strongly worsened, and the additional interactions established by the alcoholic R_2 group do not compensate for the loss of interaction strength of the zinc-binding group. Apparently, a sizable steric misfit is being faced in this derivative. Figure 2C shows that there is almost no perturbation of the hydrophobic interactions of the biphenyl group with respect to those experienced by the same group in the R_2 = -H derivative. This suggests that the two phenyl rings together are held so tightly in the S_1' pocket that the steric perturbation introduced by the R_2 substituent on the sulfonamide nitrogen, which is apparently tolerated in the R_1 = -F and R_1 = -OCH_3 derivatives, is not tolerated by the R_1 = -C_6H_5 derivative. As a consequence, a serious mismatch of the hydroxamic moiety in the zinc subsite occurs. It should be recalled that the hydroxamic group, if free to optimize its orientation, is very specifically tailored for the MMP active site, featuring bidentate binding to zinc together with two strong hydrogen bonds, one with the carboxylate side chain of the active site Glu-219 and one with the peptide carbonyl group of Ala-182.

R_3 = *-D,L-Ser* or *-D,L-Thr*. The effect of these substituents (**8**-**11**) has been examined on the scaffold having R_1 = -OCH_3. The D stereoisomers have higher affinity, and the L stereoisomers lower affinity for MMP12 with respect to the R_3 = -H scaffold (Figure 7). Examination of Figure 4A-D again provides a full rationale for this behavior: The D-isomers, besides providing additional binding interactions with the R_3 group, also exert a slight stabilizing effect on the hydroxamic moiety. The stabilization is somewhat larger for the D-Thr derivative, but is accompanied by a

modest weakening of the sulfonamide nitrogen interactions, so that the overall gain in affinity is slightly lower for the D-Thr (**9**) than for the D-Ser (**8**) derivative. In the L-Ser stereoisomer (**10**) the stabilizing effect on the hydroxamic moiety is estimated to be essentially lost, leading to a very small overall stabilization effect, while the experimental $\Delta\Delta G^0$ indicates an overall modest destabilization. Finally, the behavior of the R_3 = L-Thr (**11**) is strikingly different, both the experiment and the computational estimate showing a strong destabilization. Figure 5A-B shows that the L-Thr group actually has slightly more stabilizing interactions that the D-Thr group; however, its stereochemistry causes an appreciable distortion of the scaffold that apparently strongly destabilizes *both* the hydroxamic moiety and the sulfonamide NH and SO_2 groups. Again, it is remarkable that the experimental $\Delta\Delta G^0$ trend is reproduced so well by these simple calculations (Figure 7).

Figure 7 also shows the direct differences between the L and D isomers, and confirm the analysis. In the case of the Ser derivatives (**8**, **10**) the L isomer slightly stabilizes the Ser itself and mainly destabilizes the hydroxamic moiety, while in the case of the Thr derivatives (**9**, **11**) the Thr itself is slightly stabilized but the hydroxamic moiety and the sulfonamide NH and SO_2 groups are substantially destabilized. This direct comparison between the L and D isomers is of course redundant, because the effects could have been deduced from the preceding panels of Figure 7. However, it is instructive because the $\Delta\Delta G^0$ values in this latter comparison are by definition free from solvent-ligand interactions and essentially free also from solvent-protein interactions. In other words, the $\Delta\Delta G^0$ values almost purely reflect ligand-protein interactions.

R_2 = -$CH_2CH(CH_3)_2$; R_2 = -$CH_2CHOHCH_2OH$ with R_3 = -H, -CH_2OH. Having established that the -OCH_3 substitution in R_1 allows substituents in either R_2 or R_3 to increase the overall affinity for MMP12, the relative effects of hydrophobic vs. hydrophilic substituents in R_2 was investigated for R_1 = -OCH_3, both in the absence (-H) or in the presence (-CH_2OH) of a substituent in R_3 with a D configuration. The experimental and predicted $\Delta\Delta G^0$ values for hydrophobic (**12**) vs. hydrophilic (**13**) R_2

are shown in Figure 7, and the calculated energetics of the individual interactions are reported in Figure 6A-D. The estimated sign of the effect is again correct, *i.e.* both R_2 groups are predicted and found to stabilize the R_1 = -OCH_3, R_2 = -H, R_3 = -H scaffold, although the stabilizing effect is somewhat overestimated for the hydrophobic substituent and underestimated for the hydrophilic substituent. In the first case, Figure 6A shows that the hydrophobic substituent has by itself a strong stabilizing effect, on top of which a non-negligible stabilization of the hydroxamic moiety is also observed. In the second case (Figure 6B), the stabilizing effect of the R_2 substituent is smaller, but the stabilizing effect on the hydroxamic moiety is maintained. These findings are consistent with what observed for the R_2 = -CH_2CH_2OH case already discussed: it can be concluded that for R_1 = -OCH_3 (but not R_1 = -F or R_1 = -C_6H_5) any substituent at R_2 improves the binding of the hydroxamic group. With this in mind, the derivative with R_2 = -$CH_2CHOHCH_2OH$ and R_3 = -CH_2OH simultaneously present (**14**) was tested. These R_2 and R_3 groups, taken separately, both improved the binding of the starting R_1 = -OCH_3 scaffold. Figure 7 shows that the effect is not synergistic, the improvement being of the same order of each substituent separately, and actually slightly smaller than that caused by the R_3 = -CH_2OH substituent alone. Again, the calculations are in qualitative agreement in all cases. Figure 6C-D surprisingly shows that the hydroxamic moiety and the SO_2 groups are both improved by the simultaneous presence of the two substituents, but each substituent adopts a less optimal set of interactions with the protein, essentially neutralizing the improvement.

Implications for drug-design strategies. The above analysis demonstrates that the availability of high resolution structures of a series of enzyme adducts with homologous ligands provides precious information on the subtle factors that modulate ligand affinity, and thus may be of much help in the optimization of such ligands. For instance, increasing the size of the hydrophobic group in the S_1' pocket increases the affinity, but only provided that no substituents are placed in R_2 and R_3 positions. Conversely, a smaller hydrophobic group such as the one with R_1 = -OCH_3

permits the insertion of other groups, including more hydrophilic groups, in R_2 and R_3, provided their local effect is not destabilizing. One could even conceive substitutions in R_1, R_2 and R_3 that strongly destabilize the hydroxamic moiety but provide enough local stabilization to partially compensate for the loss. Under such circumstances, the substitution of the now unfit hydroxamic moiety with a more druggable group with a better fit could provide enough extra stabilization to obtain nanomolar affinity for a non-hydroxamic ligand. Regarding the R_2 and R_3 substituents, the present results also show that hydrophilic substitutions are not severely disfavored with respect to hydrophobic substitutions, as commonly assumed for this class of molecules. Furthermore, as other MMPs such as MMP13 have a larger binding site, one could exploit the present findings to test ligands with substituents in *e.g.* R_1, R_2 that show more or less pronounced steric misfit in MMP12 for improved selectivity for MMP13. Indeed, ΔG^0 data on binding of selected ligands to MMP13 confirm the correctness of the approach (Figure 8): adding an R_2 = CH_2CH_2OH to the R_1 = OCH_3 or R_1 = C_6H_5 scaffolds stabilizes the adduct with MMP13 more, or destabilizes it less, than the corresponding adduct with MMP12, resulting in both cases in an increased selectivity for MMP13. Of course, if suitable crystals for ligand soaking were available for other MMPs, the potential of this strategy would be further enhanced. Finally, The degree of prediction achieved with the present strategy should help avoiding the extensive screening of substituents in *n* different positions (R_1, R_2 and R_3 in this case) by examining a full *n*-dimensional matrix of compounds. Indeed, once the interplay among a certain set of substituents shows that one position is disfavored, examining further substitutions in that position with the chosen set of substituents can be avoided.

It should be stressed that the above findings have only been made possible by the availability of experimental structures. While docking programs are able to provide reasonable models of ligand binding modes, especially in the presence of existing structures with homologous ligands, the accuracy of the resulting models is lower than that of experimental structures, and insufficient for the present purpose. This has

been tested by taking the high-resolution (1.1 Å) crystal structure of MMP12 and using it to dock all the inhibitors described in this work. Despite the overall binding mode always resulted correct, the details of the interatomic distances were not accurate enough to make any energetic predictions. One drawback is the use of the same protein structure for all adducts, while the present work shows that the protein structures differ slightly from one another in the various adducts, as a consequence of some mutual adaptation or "induced fit". Attempts to mimic this induced fit, either by allowing protein side-chain movements during docking or by overall (*i.e.* ligand and protein) restrained energy minimization of the final adduct did not improve the energetic predictions significantly (data not shown). On the other hand, it is reassuring that the choice of the set of potentials to extract energetic information from the experimental structures is not critical, as different popular potential sets provide similar results.

Thermodynamics of ligand binding. Of course, it is well known that the relationship between ΔG^0 and intermolecular potential energy calculations is totally empirical, although it is often found to hold.[45-47] There are two reasons to question its applicability. The first reason is that contributions to ΔG^0 from solvent-ligand, solvent-protein and solvent-adduct interactions are not taken into account. However, it can be speculated that, at least in systems of this kind, solvation effects are not so strong, and there is always partial cancellation between solvent effects on the reactants and the product. Furthermore, in our approach $\Delta\Delta G^0$ values are estimated, so the differential solvent-protein effects are zero, and the other two terms are likely to be further reduced. In one case, *i.e.* the comparison between the behavior of the D- and L- isomers substituted in R_2, solvent effects on the ligands are also zero, and the differential solvent-adduct effects are probably negligibly small. It is obvious that the striking difference in binding affinity between the R_2 = D-Thr and R_2 = L-Thr is almost entirely due to the difference in protein-ligand interaction.

The second reason is that, even if the leading term is the free energy of the protein-ligand interaction, the latter is still composed by enthalpic and entropic terms,

and intermolecular interaction energy calculations only take into account the enthalpic term (except for a modest correction for possible differences in number of immobilized rotatable bonds). Prompted by these considerations, ΔH^0 values for the binding of some of the present ligands were measured through isothermal titration calorimetry (ITC). Two other well known strong MMP ligands, actinonin and galardin (Biomol Inc.), were also studied for comparison purposes.

ITC measurements. Isothermal titration calorimetry is widely used in drug discovery, since it permits an accurate and extensive thermodynamic characterization of protein-ligand interactions.[48,49] In particular, ITC measurements have already been successfully exploited to design and characterize Stromelysin inhibitors.[50,51]

All measurements were performed at 298 K. At variance with fluorimetric assays, here the protein concentrations were much higher (20 μM), and the AHA concentration was correspondingly higher (4 mM). This concentration is comparable with the dissociation constant of AHA itself. Therefore, AHA interferes with the binding of the inhibitors, and the experimental ΔH_0 values are conditional values. A proper estimate of the ΔH_0 for AHA was needed, both to obtain the true ΔH_0 values for the other inhibitors and to obtain thermodynamic parameters of AHA itself. A series of ITC measurements on the binding of the **3** (R_1 = OCH_3, R_2 = H) were thus performed in the presence of various AHA concentrations in the range 1 to 256 mM. By fitting the obtained enthalpy as a function of AHA concentration, a ΔH^0 of –3.180 kcal/mol was found for the binding of AHA. The latter parameter provides an experimental estimate of the enthalpic contribution of the metal binding group to the overall binding of all the inhibitors investigated here, as well as of all hydroxamic-based inhibitors in general, to MMP12. Furthermore, the AHA ΔH^0 value allowed us to estimate the correction to apply to the other ΔH^0 values. The corrected thermodynamic parameters of some of the inhibitors investigated here, as well as those of AHA, are reported in Table 1.

Although ITC measurements performed at increasing inhibitor concentrations as in the present experimental scheme also provide estimates of the dissociation

constants, the latter are affected by large errors when the constants approach nanomolar values, and so are the derived ΔG^0 values. Conversely, the ΔH^0 values are not affected by such errors. 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. The $-T\Delta S^0$ values are also reported in Table 1. The binding constants derived from ITC measurements (not shown) were in any case within a factor two with respect to the fluorimetric values.

From the the ΔH^0 and $-T\Delta S^0$ values reported in Table 1 it appears that the free energy of binding of all the homologous ligands examined in this work contains an important enthalpic contribution, and a less important but sizable entropic contribution, both favorable to binding. This appears a general feature for this class of ligands, but is by no means general for other inhibitors of MMPs. For instance, galardin and actinonin, two well known strong inhibitors of MMPs, show much less favorable, or even unfavorable enthalpic contributions, respectively, and much more favorable entropic terms (Table 1).

By examining the ΔH^0 values in Table 1 no apparent correlation can be seen with the corresponding ΔG^0 values, neither with the intermolecular interaction energies estimated from the experimental structures in the present homologous series of ligands; nor do the differences $\Delta\Delta H^0$ correlate better with the $\Delta\Delta G^0$. Apparently, the enthalpies (and entropies) of the ligand-solvent interactions are sizable and vary in a rather unpredictable way, while they largely cancel each other to give a modest contribution to ΔG^0. Apparently, this is yet another manifestation of the well-known entropy-enthalpy cancellation phenomenon.

An interesting finding from ITC measurements is the fact that, as already observed for Stromelysin[51] the binding of AHA to MMP12 (and thus presumably to all others MMPs) is almost exclusively enthalpy-driven. The ΔH^0 value of about -3.2 kcal/mol for this ligand having millimolar affinity is only two-three times smaller than the binding enthalpy of nanomolar affinity ligands, which is a relevant contribution. The $-T\Delta S$ value for AHA is close to zero. This relatively strong

contribution to the overall binding enthalpy could be even higher if the ligand solvent contribution were not unfavorable, as it is expected to be for the very soluble AHA molecule. Indeed, in our parameterization, the estimated contribution of the AHA moiety to the overall protein-ligand interaction energy has to be as high as -6 kcal/mol if a reasonable estimate of the overall interaction energy for this class of ligands has to be obtained. This is not unespected, as it is well known that, if AHA is considered a fragment of a strong binding ligand rather than an independent molecule, the loss of rigid body rotational and traslational entropy upon binding occurs only once and not twice as it would be for the isolated fragments. The corresponding gain in ΔG^0, which is parametrically incorporated in the interaction energy in our calculations, can easily be of the order of –3 kcal/mol.[51-53]

Concluding Remarks

In summary, it is shown here that an analysis of the details of the protein-ligand interactions – as they appear from the three-dimensional structures – in terms of simple hydrophobic and electrostatic contributions provides a remarkable semi quantitative account of i) the ΔG^0 values of the adducts, ii) the differences in ΔG^0 ($\Delta\Delta G^0$) between pairs of closely related ligands, and iii) the identity of the individual ligand atom-protein atom interactions that contribute in each case to increase – or decrease – the affinity of each particular ligand with respect to its closest analog. This work demonstrates that this kind of rationalization is possible and reliable, and provides hints for the planning of finely tuned inhibitors, originating from the same scaffold, for any structurally well-characterized biological target.

For some of the present ligands, the enthalpies of binding (ΔH^0) were also measured through isothermal calorimetry. All the ligands from this class are found to be characterized by favorable enthalpic as well as entropic contributions. The enthalpic contribution of the hydroxamic moiety, a popular zinc-binding group in the MMP inhibitor landscape, has been separately evaluated and discussed in terms of its relative contribution to the potency of hydroxamic-based MMP ligands. A detailed

analysis of the ΔH^0 and ΔS^0 values, taken separately, is less instructive, due to the well-known entropy-enthalpy compensation effects.[48] As in many other thermodynamic studies, the two terms are found to provide opposite and largely canceling contributions to ΔG^0.

Acknowledgements. This work was supported by EC (Projects: LSHG-CT-2004-512077and TOC: MTKI-CT-2004-509750), MIUR (PRIN 2005, Prot. N. 2005039878, Prot. N. 4455, and RBAU013NSB), and Ente Cassa di Risparmio di Firenze (Innovative Strategies for Drug Design).

References

(1) Fisher, J. F.; Mobashery, S. *Cancer Metastasis Rev.* **2006,** *25,* 115-136.

(2) Whittaker, M.; Floyd, C. D.; Brown, P.; Gearing, A. J. *Chem. Rev.* **1999,** *99,* 2735-2776.

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

(4) Xue, C. B.; He, X. H.; Corbett, R. L.; Roderick, J.; Wasserman, Z. R.; Liu, R. Q.; Jaffee, B. D.; Covington, M. B.; Qian, M. X.; Trzaskos, J. M.; Newton, R. C.; Magolda, R. L.; Wexler, R. R.; Decicco, C. P. *J. Med. Chem.* **2001,** *44,* 3351-3354.

(5) 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.

(6) Hajduk, P. J.; Sheppard, G.; Nettesheim, D. G.; Olejniczak, E. T.; Shuker, S. B.; Meadows, R. P.; Steinman, D. H.; Carrera, G. M.; Marcotte, P. A.; Severin, J.; Walter, K.; Smith, H.; Gubbins, E.; Simmer, R.; Holzman, T. F.; Morgan, D. W.; Davidsen, S. K.; Summers, J. B.; Fesik, S. W. *J. Am. Chem. Soc.* **1997,** *119,* 5818-5827.

(7) MacPherson, L. J.; Bayburt, E. K.; Capparelli, M. P.; Carroll, B. J.; Goldstein, R.; Justice, M. R.; Zhu, L.; Hu, S.; Melton, R. A.; Fryer, L.; Goldberg, R. L.; Doughty, J. R.; Spirito, S.; Blancuzzi, V.; Wilson, D.; O'Byrne, E. M.; Ganu, V.; Parker, D. T. *J. Med. Chem.* **1997,** *40,* 2525-2532.

(8) Coussens, L. M.; Fingleton, B.; Matrisian, L. M. *Science* **2002,** *295,* 2387-2392.

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

(10) Zartler, E. R.; Shapiro, M. J. *Curr. Opin. Chem. Biol.* **2005,** *9,* 366-370.

(11) Erlanson, D. A.; Wells, J. A.; Braisted, A. C. *Annu. Rev. Biophys. Biomol. Struct.* **2004,** *33,* 199-223.

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

(13) Moy, F. J.; Chanda, P. K.; Chen, J. M.; Cosmi, S.; Edris, W.; Levin, J. I.; Powers, R. *J. Mol. Biol.* **2000,** *302,* 671-689.

(14) Brandstetter, H.; Grams, F.; Glitz, D.; Lang, A.; Huber, R.; Bode, W.; Krell, H. W.; Engh, R. A. *J. Biol. Chem.* **2001,** *276,* 17405-17412.

(15) Gall, A. L.; Ruff, M.; Kannan, R.; Cuniasse, P.; Yiotakis, A.; Dive, V.; Rio, M. C.; Basset, P.; Moras, D. *J. Mol. Biol.* **2001,** *307,* 577-586.

(16) Borkakoti, N.; Winkler, F. K.; Williams, D. H.; Darcy, A.; Broadhurst, M. J.; Brown, P. A.; Johnson, W. H.; Murray, E. J. *Nat. Struct. Biol.* **1994,** *1,* 106-110.

(17) 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.

(18) Lang, R.; Kocourek, A.; Braun, M.; Tschesche, H.; Huber, R.; Bode, W.; Maskos, K. *J. Mol. Biol.* **2001,** *312,* 731-742.

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

(20) Vandoren, S. R.; Kurochkin, A. V.; Hu, W. D.; Ye, Q. Z.; Johnson, L. L.; Hupe, D. J.; Zuiderweg, E. R. P. *Protein Sci.* **1995,** *4,* 2487-2498.

(21) Browner, M. F.; Smith, W. W.; Castelhano, A. L. *Biochemistry* **1995,** *34,* 6602-6610.

(22) Rowsell, S.; Hawtin, P.; Minshull, C. A.; Jepson, H.; Brockbank, S. M. V.; Barratt, D. G.; Slater, A. M.; McPheat, W. L.; Waterson, D.; Henney, A. M.; Pauptit, R. A. *J. Mol. Biol.* **2002,** *319,* 173-181.

(23) Rush, T. S.; Powers, R. The Application of X-Ray, NMR, and Molecular Modeling in the Design of MMP Inhibitors. *Curr. Top. Med. Chem.* **2004,** *4,* 1311-1327.

(24) Moy, F. J.; Pisano, M. R.; Chanda, P. K.; Urbano, C.; Killar, L. M.; Sung, M. L.; Powers, R. *J. Biomol. NMR* **1997,** *10,* 9-19.

(25) Gao, G.; Semenchenko, V.; Arumugam, S.; Van Doren, S. R. *J. Mol. Biol.* **2000,** *301,* 537-552.

(26) Teague, S. J. *Nature Rev. Drug Discov.* **2003,** *2,* 527-541.

(27) Moy, F. J.; Chanda, P. K.; Chen, J.; Cosmi, S.; Edris, W.; Levin, J. I.; Rush, T. S.; Wilhelm, J.; Powers, R. *J. Am. Chem. Soc.* **2002,** *124,* 12658-12659.

(28) Morris, D. G.; Huang, X. Z.; Kaminski, N.; Wang, Y. N.; Shapiro, S. D.; Dolganov, G.; Glick, A.; Sheppard, D. *Nature* **2003,** *422,* 169-173.

(29) Vos, C. M. P.; van Haastert, E. S.; de Groot, C. J. A.; van der Valk, P.; de Vries, H. E. Matrix Metalloproteinase-12 *J. Neuroimmunol.* **2003,** *138,* 106-114.

(30) Leslie, A. G. W. In Crystallographic Computing V. In *Molecular Data Processing*; Moras, D., Podjarny, A. D., Thierry, J.-C., Eds.; Oxford University Press: Oxford, **1991**.

(31) Evans, P. R. "Data Reduction", Proceedings of CCP4 Study Weekend. Data Collection & Processing, 114-122. **1993**.

(32) Rossmann, M. G.; Blow, D. M. *Acta Cryst.* **1962,** *D15,* 24-31.

(33) Crowther, R. A. The Molecular Replacement Method. Rossmann, M. G., Ed.; Gordon & Breach: New York, 1972.

(34) Vagin, A.; Teplyakov, A. *J. App. Cryst.* **1997,** *30,* 1022-1025.

(35) Vagin, A.; Teplyakov, A. *Acta Crystallogr. D Biol. Crystallogr.* **2000,** *56 Pt 12,* 1622-1624.

(36) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Cryst.* **1997,** *D53,* 240-255.

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

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

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

(40) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comp. Chem.* **1998,** *19,* 1639-1662.

(41) Hu, X.; Shelver, W. H. *J. Mol. Graph. Model.* **2003,** *22,* 115-126.

(42) Shah, A. V.; Walters, W. P.; Shah, R.; Dolata, D. P. Babel - A molecular Structure Information Interchange Hub. Lysakowski, R. and Gragg, C. E. 1994. Philadelphia, PA, American Society for Testing and Materials. **1994**.

(43) Indyk, L.; Fisher, H. F. *Methods Enzymol.* **1998,** *295,* 350-364.

(44) Hu, X.; Balaz, S.; Shelver, W. H. *J. Mol. Graph. Model.* **2004,** *22,* 293-307.

(45) Lamb, M. L.; Jorgensen, W. L. *Curr. Opin. Chem. Biol.* **1997,** *1,* 449-457.

(46) Aqvist, J.; Luzhkov, V. B.; Brandsdal, B. O. *Acc. Chem. Res.* **2002,** *35,* 358-365.

(47) Osterberg, F.; Morris, G. M.; Sanner, M. F.; Olson, A. J.; Goodsell, D. S. *Proteins Struct. Funct. Genet.* **2002,** *46,* 34-40.

(48) Holdgate, G.; Fisher, S.; Ward, W. The Application of Isothermal Titration Calorimetry to Drug Discovery. In *Biocalorimetry 2*; Ladbury, J. E., Doyle, M. L., Eds.; John Wiley & Sons Ltd: Chichester, **2004**.

(49) Ruben, A. J.; Kiso, Y.; Freire, E. *Chem. Biol. Drug Des.* **2006,** *67,* 2-4.

(50) Parker, M. H.; Lunney, E. A.; Ortwine, D. F.; Pavlovsky, A. G.; Humblet, C.; Brouillette, C. G. *Biochemistry* **1999,** *38,* 13592-13601.

(51) Olejniczak, E. T.; Hajduk, P. J.; Marcotte, P. A.; Nettesheim, D. G.; Meadows, R. P.; Edalji, R.; Holzman, T. F.; Fesik, S. W. *J. Am. Chem. Soc.* **1997,** *119,* 5828-5832.

(52) Jencks, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1981,** *78,* 4046-4050.

(53) Murray, C. W.; Verdonk, M. L. *J. Comput. Aided Mol. Des.* **2002,** *16,* 741-753.

Captions to the Figures

Figure 1. (A) Schematic representation of the active site of MMPs highlighting the zinc ion site, the hydrophobic pocket termed S_1', and the substrate binding groove. (B) The sulfonamide scaffold used in this work with its R_1, R_2 and R_3 substituents. The actual compounds investigated are:

1 R_1 = H, R_2 = H, R_3 = H

2 R_1 = F, R_2 = H, R_3 = H

3 R_1 = OCH_3, R_2 = H, R_3 = H

4 R_1 = C_6H_5, R_2 = H, R_3 = H

5 R_1 = F, R_2 = CH_2CH_2OH, R_3 = H

6 R_1 = OCH_3, R_2 = CH_2CH_2OH, R_3 = H

7 R_1 = C_6H_5, R_2 = CH_2CH_2OH, R_3 = H

8 R_1 = OCH_3, R_2 = H, R_3 = (D)CH_2OH

9 R_1 = OCH_3, R_2 = H, R_3 = (D)$CHCH_3OH$

10 R_1 = OCH_3, R_2 = H, R_3 = (L)CH_2OH

11 R_1 = OCH_3, R_2 = H, R_3 = (L)$CHCH_3OH$

12 R_1 = OCH_3, R_2 = $CH_2CH(CH_3)_2$, R_3 = H

13 R_1 = OCH_3, R_2 = $CH_2CHOHCH_2OH$, R_3 = H

14 R_1 = OCH_3, R_2 = $CH_2CHOHCH_2OH$, R_3 = (D)CH_2OH

Figure 2. Close up of the high resolution 3D structures of the adducts of MMP12 with ligands **2**-**4** (left hand column) and **1** (central column, panels A-C). The right hand column highlights the improvement (green spheres) and worsening (red spheres) of the protein interaction with individual atoms of ligands **2**-**4** with respect to **1**. The volumes of the spheres are proportional to the size of the interaction energy variation.

Figure 3. Close up of the high resolution 3D structures of the adducts of MMP12 with ligands **5**-**7** (left hand column panels A-C) and **2**-**4** (central column, panels A-C). The right hand column highlights the improvement (green spheres) and worsening (red spheres) of the protein interaction with individual atoms of ligands **5-7** with respect to **2-4**, respectively. The volumes of the spheres are proportional to the size of the interaction energy variation.

Figure 4. Close up of the high resolution 3D structures of the adducts of MMP12 with ligands **8**-**11** (left hand column panels A-D) and **3** (central column, panels A-D). The right hand column highlights the improvement (green spheres) and worsening (red spheres) of the protein interaction with individual atoms of ligands **8-11** with respect to **3**. The volumes of the spheres are proportional to the size of the interaction energy variation.

Figure 5. Close up of the high resolution 3D structures of the adducts of MMP12 with ligands **10** and **11** (left hand column panels A,B) and **8**-**9** (central column, panels A,B). The right hand column highlights the improvement (green spheres) and worsening (red spheres) of the protein interaction with individual atoms of ligands **10** and **11** with respect to **8**-**9**. The volumes of the spheres are proportional to the size of the interaction energy variation.

Figure 6. Close up of the high resolution 3D structures of the adducts of MMP12 with ligands **12**-**14** (left hand column panels A-D), and **3** (central column, panels A-C), and **8** (central column, panels D). The right hand column highlights the improvement (green spheres) and worsening (red spheres) of the protein interaction with individual atoms of ligands **12-14** with respect to **3** and **14** with respect to **8**. The volumes of the spheres are proportional to the size of the interaction energy variation.

Figure 7. Calculated (white bars) and observed (black bars) free energy variations ($\Delta\Delta G^0$) on passing from one MMP12 adduct to another with a close analogue ligand. The calculated $\Delta\Delta G^0$ values are the algebraic sum of all the intermolecular interaction energies plus a entropic penalty term related to immobilization of rotatable bonds as described in Materials and Methods. The bars extending, inwards or outwards of the calculated $\Delta\Delta G^0$ values, show the results obtained by neglecting the entropic penalty. The individual intermolecular interaction energy differences are illustrated in Figure 2 for ligand pairs **2**-**1**, **3**-**1** and **4**-**1**, in Figure 3 for ligand pairs **5**-**2**, **6**-**3** and **7**-**4**, in Figure 4 for ligand pairs **8**-**3**, **9**-**3**, **10**-**3** and **11**-**3**, in Figure 5 for ligand pairs **10**-**8** and **11**-**9**, and in Figure 6 for ligand pairs **12**-**3**, **13**-**3**, **14**-**3** and **14**-**8**.

Figure 8. Experimental free energy of binding to MMP12 (up triangles) and MMP13 (filled circles) for ligands **1**, **3** and **6** (A) and **1**, **4** and **7** (B). The increased discrimination between MMP12 and MMP13 is highlighted.

Table 1. Corrected thermodynamic parameters for the investigated inhibitors.

	ΔG^0	ΔH^0	$-T\Delta S^0$
1	-9.833	-9.095	-0.738
2	-9.796	-6.772	-3.024
3	-10.503	-8.520	-1.983
4	-11.759	-8.849	-2.910
5	-10.092	-6.459	-3.633
6	-11.046	-9.653	-1.393
7	-10.243	-	-
8	-11.216	-8.391	-2.825
9	-10.857	-8.124	-2.733
10	-10.063	-	-
11	-7.775	-	-
12	-11.428	-8.243	-3.185
13	-11.046	-8.888	-2.158
14	-11.125	-8.082	-3.043
AHA	-3.011	-3.180	0.169
Galardin	-11.046	-4.073	-6.973
Actinonin	-9.526	2.200	-11.726

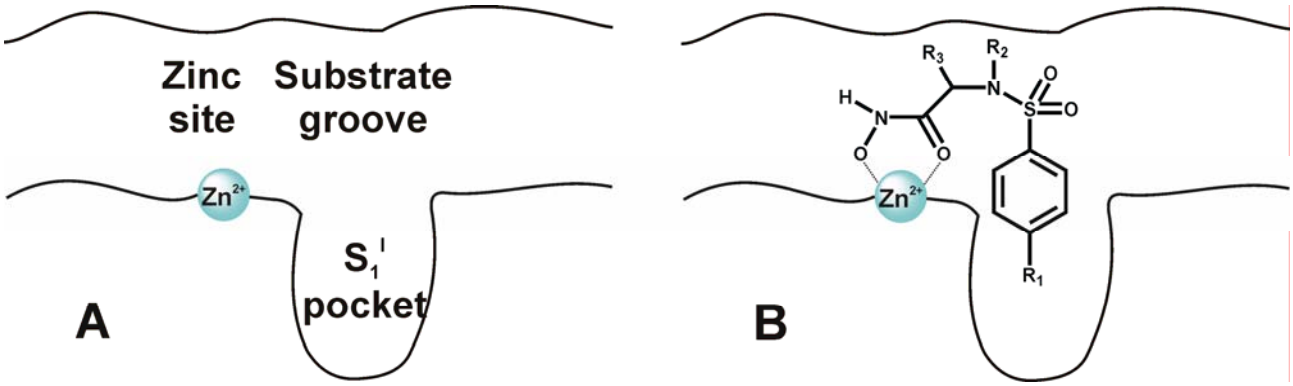

Zinc site
Substrate groove
Zn2+
S1'
pocket
A
R3
R2
H
N
N
S
O
O
O
O
O
Zn2+
R1
B

Figure 1

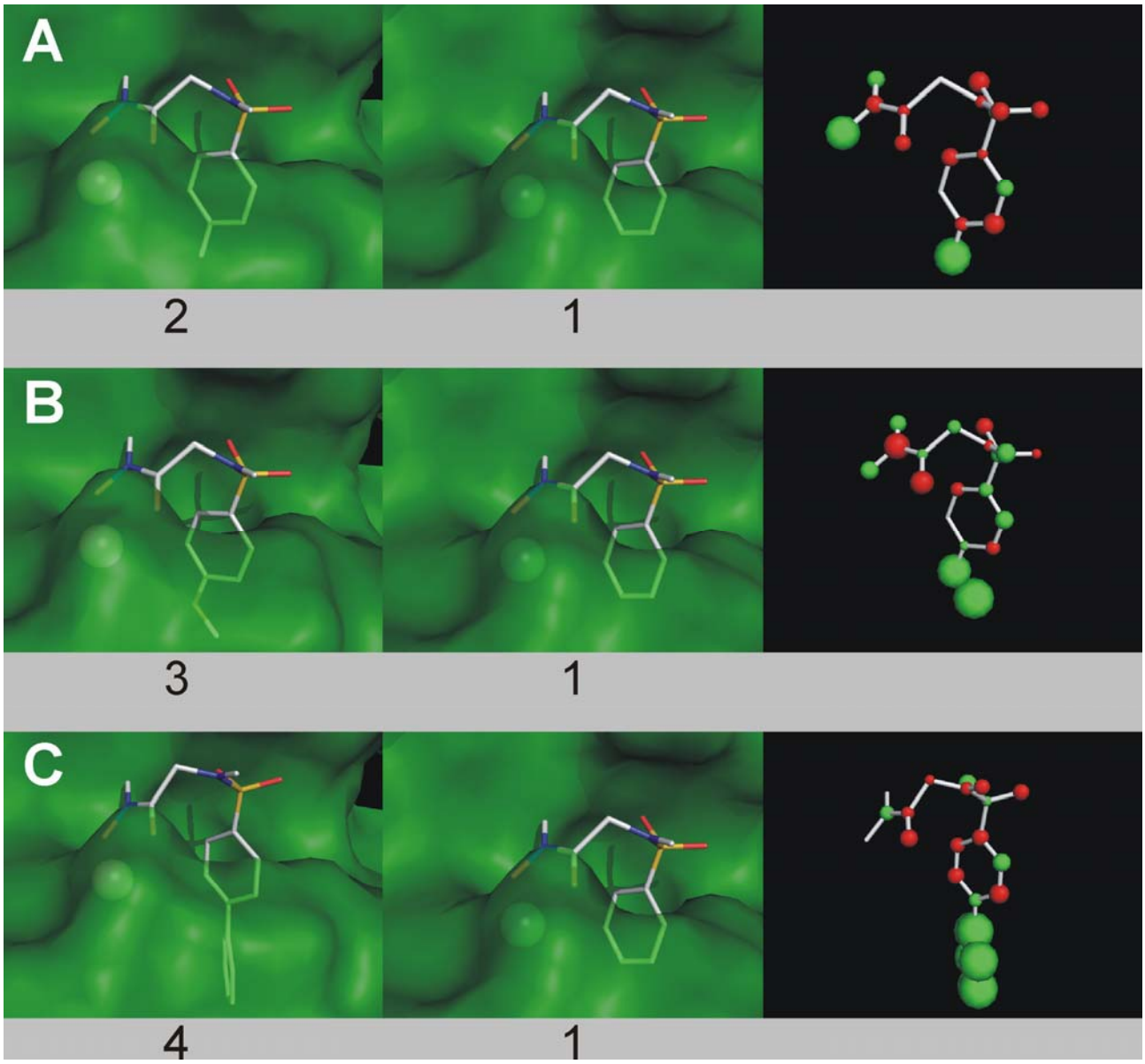


Figure 2

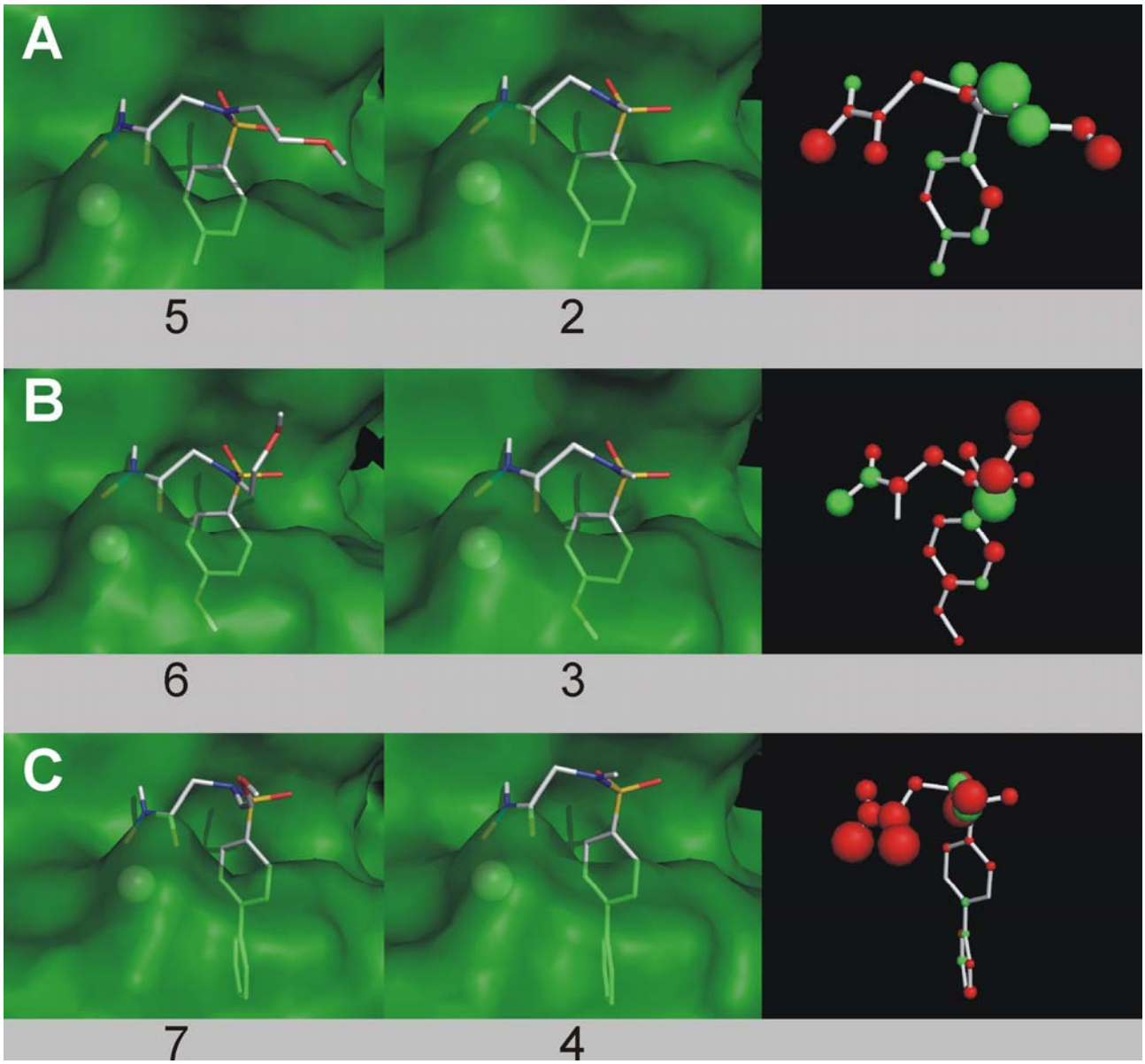


Figure 3

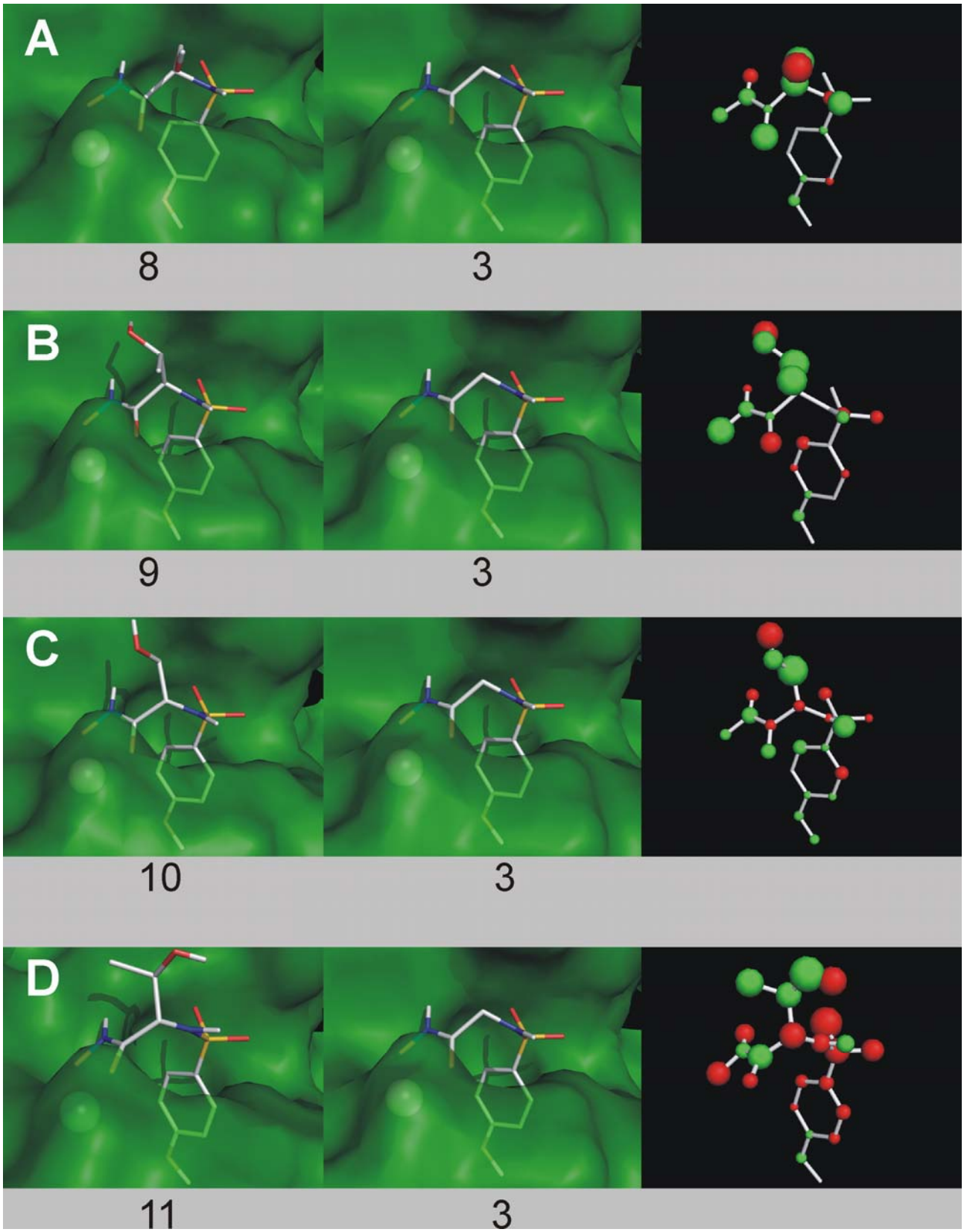


Figure 4

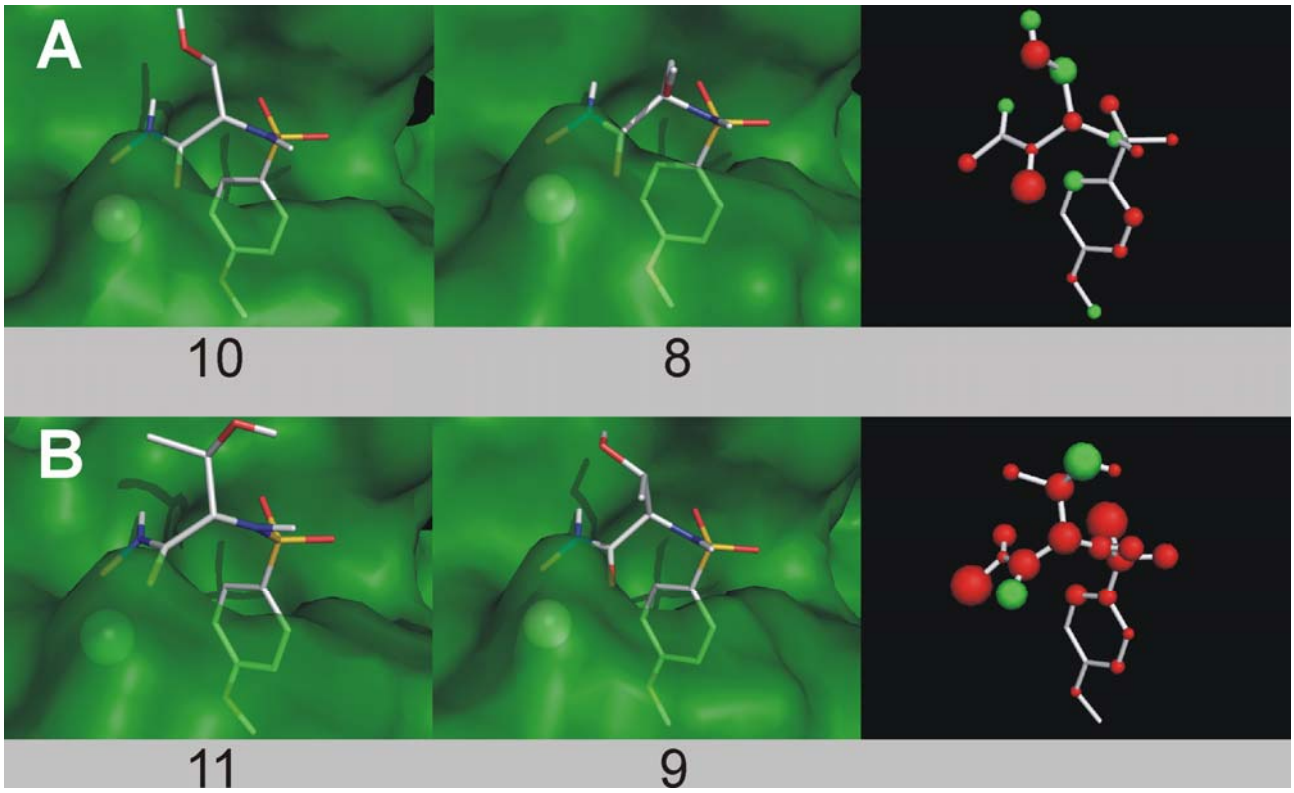


Figure 5

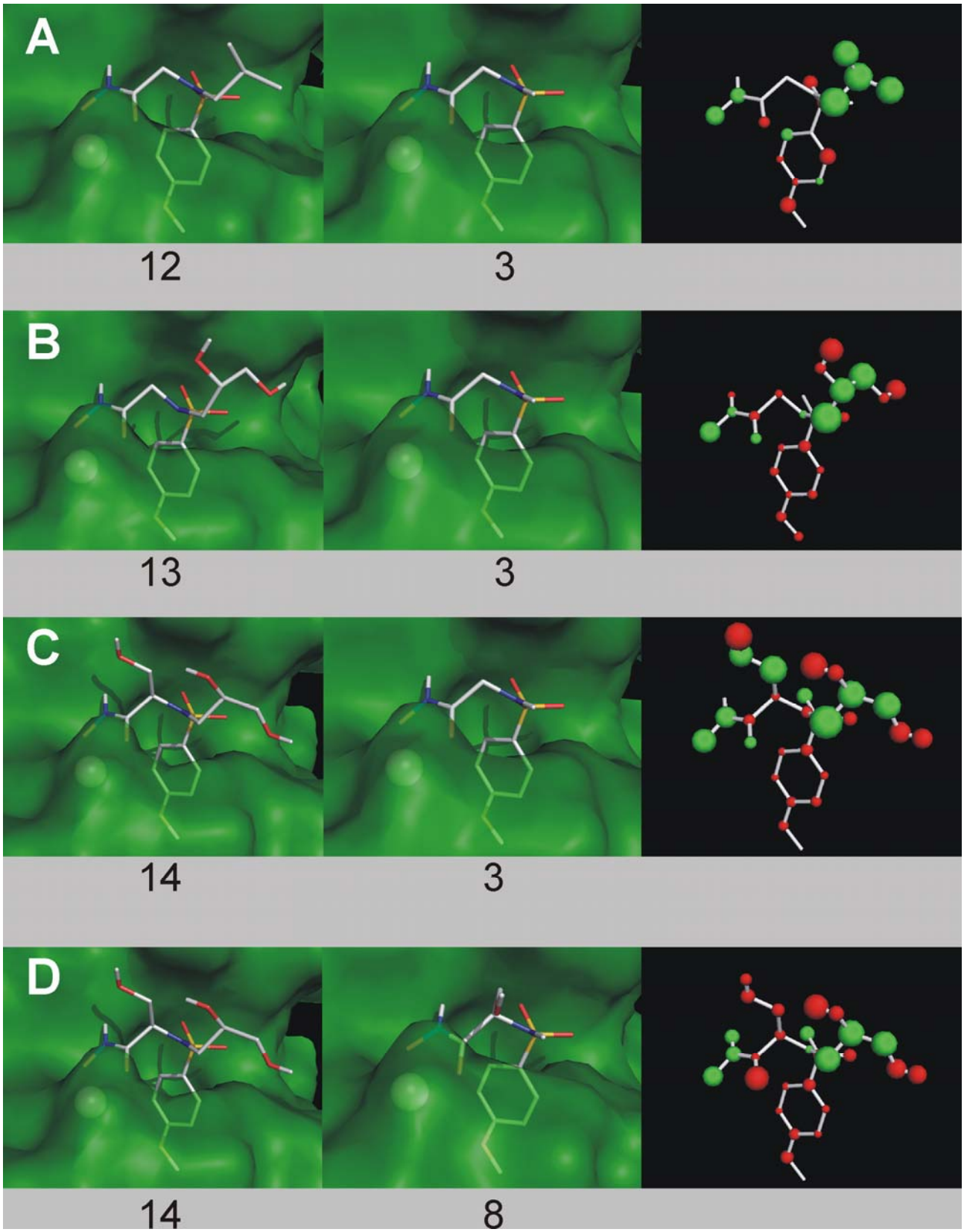


Figure 6

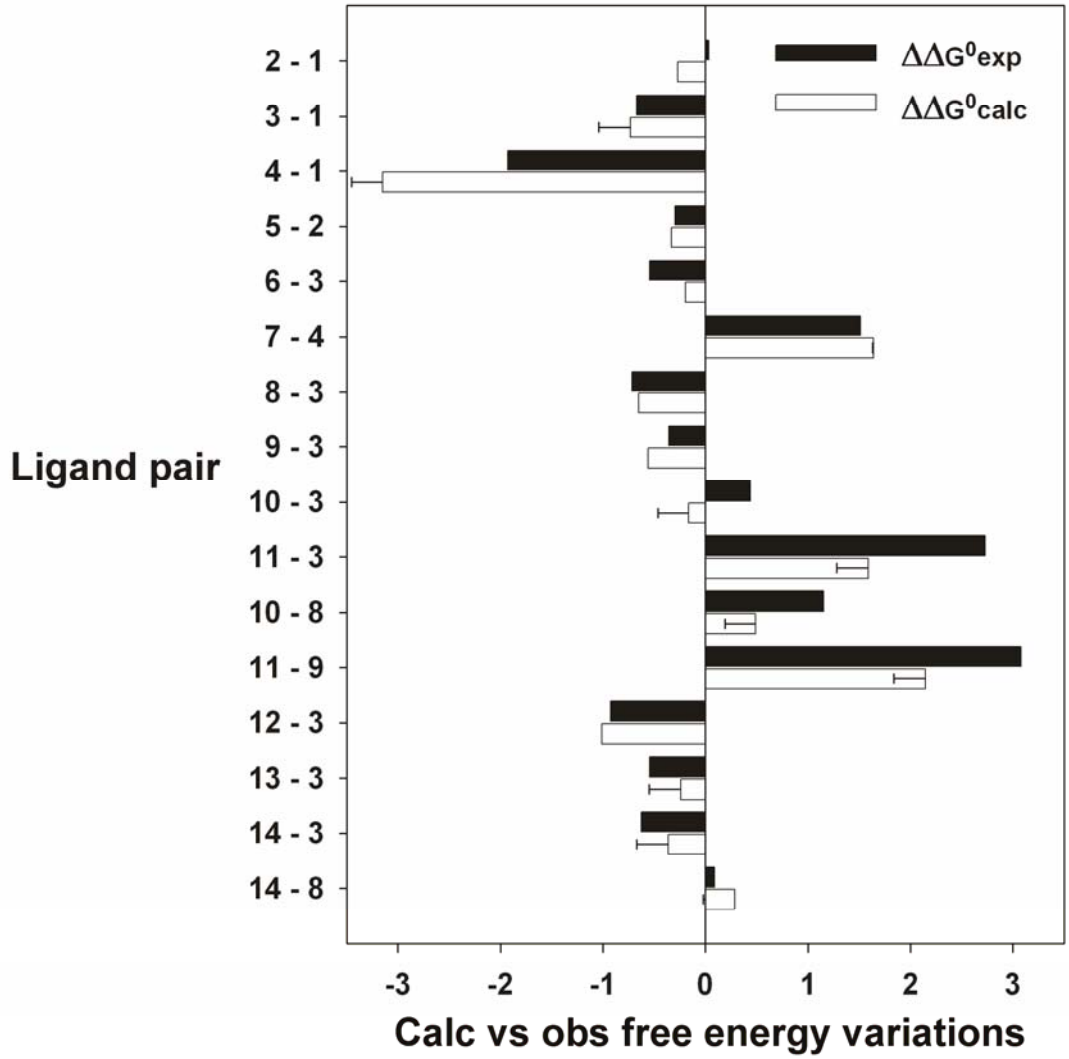

ΔΔG⁰exp
ΔΔG⁰calc
Ligand pair
2 - 1
3 - 1
4 - 1
5 - 2
6 - 3
7 - 4
8 - 3
9 - 3
10 - 3
11 - 3
10 - 8
11 - 9
12 - 3
13 - 3
14 - 3
14 - 8
-3
-2
-1
0
1
2
3
Calc vs obs free energy variations

Figure 7

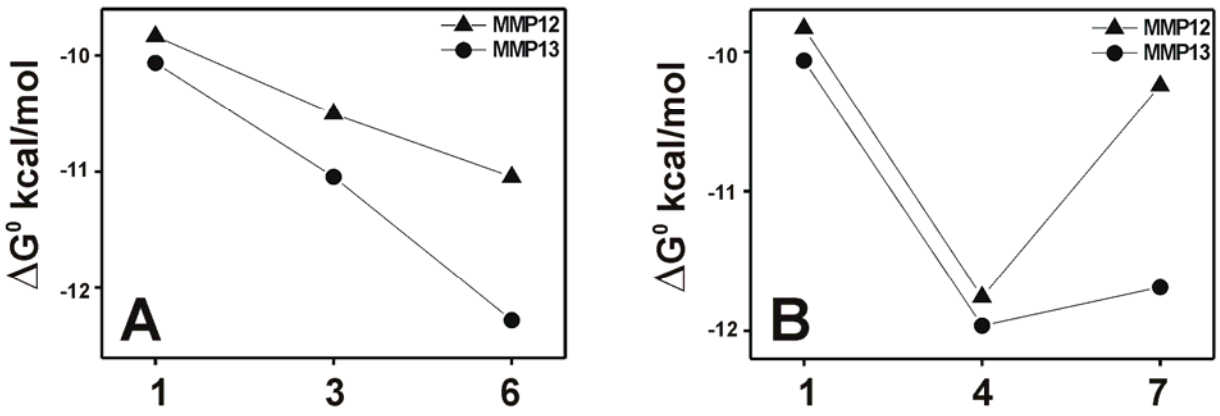

MMP12
MMP13
ΔG⁰ kcal/mol
-10
-11
-12
A
1
3
6
MMP12
MMP13
ΔG⁰ kcal/mol
-10
-11
-12
B
1
4
7

Figure 8

4.6 Solid-State NMR of Matrix Metalloproteinase 12: an Approach Complementary with Solution NMR [Accepted in *ChemBioChem*]

DOI: 10.1002/cbic.200((Please insert last 6 DOI digits))

Solid-State NMR of Matrix Metalloproteinase 12: an Approach Complementary with Solution NMR

Stéphane Balayssac,[a] Ivano Bertini,*[a],[b] Katja Fälber,[c] Marco Fragai,[a],[d] Stefan Jehle,[c] Moreno Lelli,[a] Claudio Luchinat,[a],[d] Hartmut Oschkinat,[c] and Kwon Joo Yeo[a],[b]

Solid-state NMR (SS NMR) is a technique that has shown a rapid development in recent years.[1–4] The exciting progress in sample preparation methods,[5] tailored pulse sequences,[4,6–8] and instrumentations, make now possible to investigate relatively large proteins. In spite of this development, the number of proteins for which an almost complete solid-state assignment is available is still limited.[9] Interestingly, it is generally observed that ^{13}C chemical shifts do not change much passing from the solution assignment to the one observed in microcrystalline samples (the differences are generally <1 ppm).[9] This opens the way to a fast liquid-based solid-state assignment, in which the available liquid assignment is transferred to the solid-state spectra, and only a minimal number of solid-state spectra are acquired, to unambiguously confirm the assignment obtained in solution. The possibility to have, in a short time, an attainable solid-state assignment is valuable, as it opens the possibility to investigate the same protein as part of larger aggregate (oligomerization, protein-protein complexes) without the line broadening associated to the increase in molecular weight in solution.

In the present work we demonstrate that it is possible to rapidly provide a large fraction of the solid-state NMR assignment of a relatively large protein (17 kDa), using a pair of experiments (CP MAS PDSD, *J*-dec PDSD), which can be acquired in a limited amount of time (12-15 h each), and manually assigned in few days using the available liquid-state assignment as a guideline. 3D NCACX and NCOCX PDSD spectra fully validate the assignment and further increase the overall fraction of assigned peaks, although they require considerably more experimental time.

Thus, we investigated through SS NMR a microcrystalline sample of the catalytically active domain of the zinc-containing matrix metallo-proteinase 12 (Zn-MMP-12, 159 AA, 17.6 kDa), for which the X-ray structure, and the solution NMR assignment are available.[10]

The crystallographic structure indicates that the active site domain is composed by three α-helices (44 AA, 28% of the total residues), and by seven *β*-strands (27 AA, 17%). The remaining 88 residues form unstructured turns. The C-terminal coordinated Zn plays a role in the enzymatic activity.

The SS NMR sample was prepared by crystallizing the protein from PEG 8000 30%, following the already published procedure.[10] The microcrystalline precipitate is formed just after 12 h, and the crystallization is complete after 1-2 days.

Figure 1 reports the ^{13}C–^{13}C CP MAS PDSD spectrum acquired at 16.4 T (700 MHz of ^{1}H Larmor frequency) and MAS frequency $\omega_R/2\pi$ = 11.5 kHz. Several spectra were acquired with variable mixing time of 5 and 15 ms, and MAS frequency 8.5 kHz. As it appears from Figure 1, the spectra are extremely well resolved. The assignment of the PDSD was directly obtained on the basis of the ^{13}C chemical shifts determined in solution.[10] The good resolution of the spectrum in Figure 1 allowed us to assign up to 75% of the ^{13}C aliphatic spins with minor adjustments (<1 ppm) with respect to the solution shifts (Table 1), and checking the consistency with the spin pattern of each residue as far as the aliphatic part of the spectrum is concerned. The limitation to further extend the assignment transfer arises from some ambiguities, essentially for the more crowded region of the Cα–Cβ correlations and for the leucine Cγ–Cδ correlations.

The carbonyls can be safely assigned for the 17 Gly (using the 8.5 kHz MAS frequency PDSD to avoid superposition with the C' spinning sidebands), and for other 60 residues (48% of the overall carbonyls, Table 1), while the aromatic sidechains remain unassigned.

A significant improvement in the solid-state assignment was achieved using the *J*-decoupled PDSD sequence (*J*-dec PDSD).[11] This sequence provides an enhanced resolution in the Cα–Cx correlations (with Cx = C', Cβ, Cγ, etc.) refocusing the $J_{C\alpha-C\beta}$ and $J_{C\alpha-C'}$ scalar coupling in the indirect dimension (*F1*). These couplings can also be refocused in the direct dimension if the acquired FID is properly transformed using the Maximun Entropy processing (MaxEnt). Even without MaxEnt processing (i.e. refocusing only the *J*-couplings in *F1*) the PDSD spectrum appears considerably more resolved (Figure 2) allowing one to nearly complete the assignment of the Cα–Cβ correlations (86%

[a] *Dr. S. Balayssac, Prof. I. Bertini, Dr. M. Fragai, Dr. M. Lelli, Prof. C. Luchinat, K.J. Yeo*
Magnetic Resonance Center (CERM)
University of Florence
Via Luigi Sacconi 6, 50019 Sesto Fiorentino (Italy)
[b] *Prof. I. Bertini, K.J. Yeo*
Department of Chemistry
University of Florence
Via della Lastruccia 3, 50019 Sesto Fiorentino (Italy)
[c] *Dr. K. Fälber, S. Jehle, Prof. H. Oschkinat*
Leibnizinstitut für Molekulare Pharmakologie
Robert-Rössle-Straße 10, 13125 Berlin (Germany)
[d] *Dr. M. Fragai, Prof. C. Luchinat*
Department of Agricultural Biotechnology
University of Florence
Via Maragliano 75-77, 50144 Florence (Italy).
* *Correspondence to:*
Ivano Bertini
Fax (+39)055-457-4271
E-mail: ivanobertini@cerm.unifi.it

Supporting Information, including the Experimental Part and the assignment table, is available on the WWW under http://www.chembiochem.org/ or from the author.

^{13}C aliphatic spins) and to assign up to 70% of the C' nuclei (Table 1). This assignment was fully validated using 3D NCACX PDSD and 3D NCOCX PDSD experiments.[4,12] The third dimension, introduced with the ^{15}N indirect evolution, provides an improved resolution that allowed us to resolve those ^{13}C shifts which were partly overlapped in the previous spectra (see Figure S2 and S2).

These 3D spectra, together with the previous PDSD and *J*-dec PDSD spectra, also allowed us to further increase the fraction of assigned resonances (93%, with 92% aliphatic and 93% C' nuclei, Table 1). In such a way, we arrived to assign up to 94% of the backbone and 90% of the sidechain resonances. The time required to obtain this increase is however rather long (6 days for each 3D and roughly 1-2 weeks for the spectral assignments) with respect to the 2D PDSD-based assignment. On the other hand, although this was not the primary goal of this work, we also found that the present sets of 2D and 3D spectra can provide more than 70% assignment *independently* of the availability of the solution data, which is also a remarkable result. We do not exclude that the use of other tailored experiments allow one to assign even more, resolving some ambiguities of the present 3D or assigning some residues of the N- and C- terminal part which are lost in the solution assignment probably because of the mobility in solution.

The success of the liquid-based assignment is essentially ascribable to the relatively small differences among liquid and microcrystalline solid-state chemical shifts. In Figure 3 we plot, for each residue, the differences between the solid-state and liquid–state ^{13}C chemical shifts. It is apparent how the differences are limited: the overall RMSD is 0.52 ppm, with about 93 % of the resonances that differ by less than 1 ppm. Only 6% are more than 1 ppm shifted, and only 3 resonances shift more than 2 ppm (1%). Interestingly, the deviations larger than 2 ppm belong to the resonances of Ile 180, a residue demonstrated to be involved in conformational equilibria in solution, but to be frozen in slightly different conformations in X-ray structures of the protein bound to different inhibitors.[10] We would like to stress that comparing PDSD *patterns* rather than isolated cross peaks is a very powerful way of transferring side chain assignments, which tolerates well differences in shifts of 1-2 ppm, as observed here. In this respect, we also notice that the similarity of the ^{15}N shifts between solid state and solution is lower than that observed for ^{13}C (see Figure S3), as previously observed.[13] Nevertheless, for the present system, 90% of the nitrogen resonances could be assigned by the combined approach described above.

In summary, we report the almost complete solid-state assignment of the ^{13}C (see also Figure 4), and ^{15}N nuclei of microcrystalline Zn-MMP-12, which, up to now, is the largest protein for which a solid-state NMR assignment is available. The assignment was extended up to 93% for ^{13}C and 90% for ^{15}N, and we demonstrate that it is possible to safely translate a large fraction of the ^{13}C liquid assignment to solid-state in less than a week. The possibility to have both solid-state and liquid state assignments for relatively large proteins may be an important asset to investigate biological processes at the molecular level. Indeed, many processes such as oligomerization or protein complex formation, which start from solution state, are potentially easily investigated in the solid-state NMR, which is much less affected by the line-broadening due to the increased molecular weight. Even protein-ligand interactions may be conveniently followed by SS NMR.

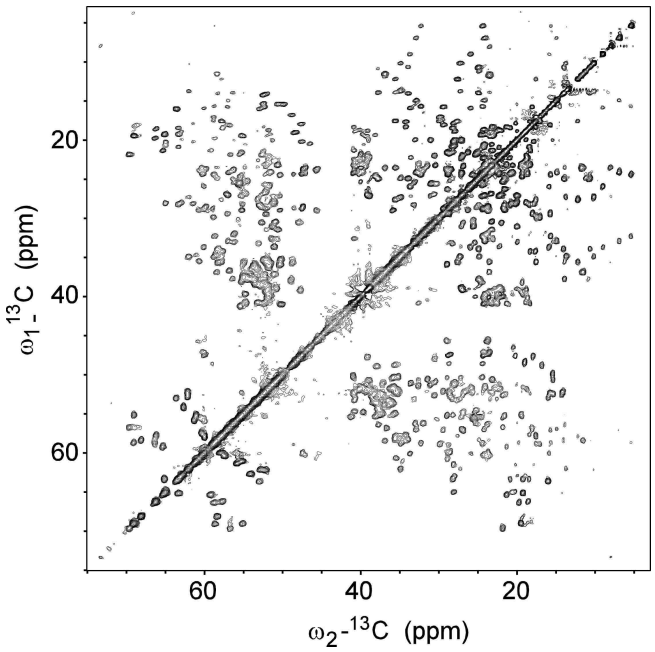


Figure 1. Aliphatic region of the ^{13}C–^{13}C CP MAS PDSD of the Zn-MMP-12 (16.4 T, 11.5 kHz MAS frequency, 15 ms mixing time).

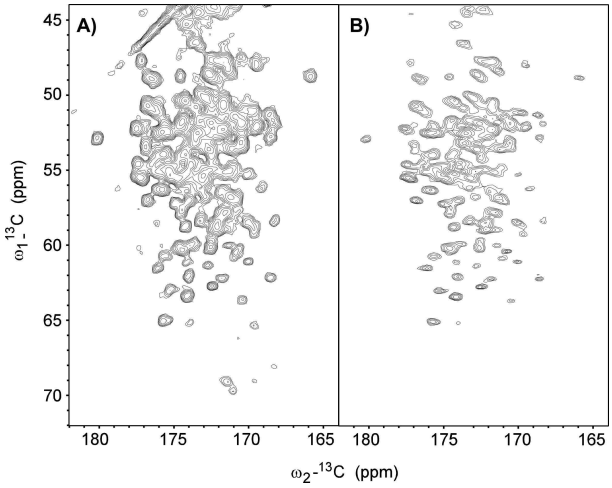


Figure 2. CACO regions in ^{13}C–^{13}C PDSD (left side) and ^{13}C–^{13}C *J*-decoupled PDSD (right side) of Zn-MMP-12 (16.4 T, 11.5 kHz MAS frequency, 15 ms mixing time for both spectra). The *J*-decoupled PDSD was processed with Fourier transformation, which retains scalar coupling in *F2*. Notice the resolution improvement with respect to the fully coupled PDSD.

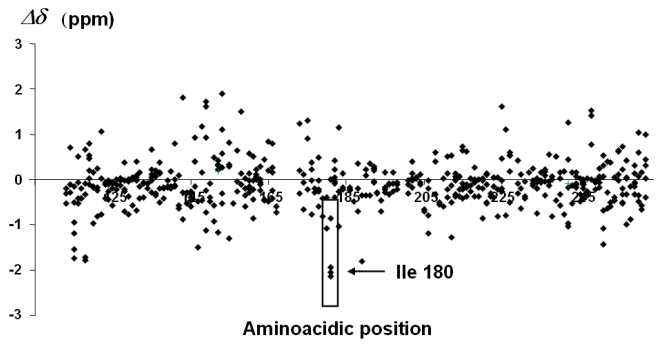


Figure 3. Plot of the ^{13}C chemical shift differences among solid-state (δ_S) and liquid state (δ_L) chemical shifts ($\Delta\delta = \delta_S - \delta_L$) for each residue positions.

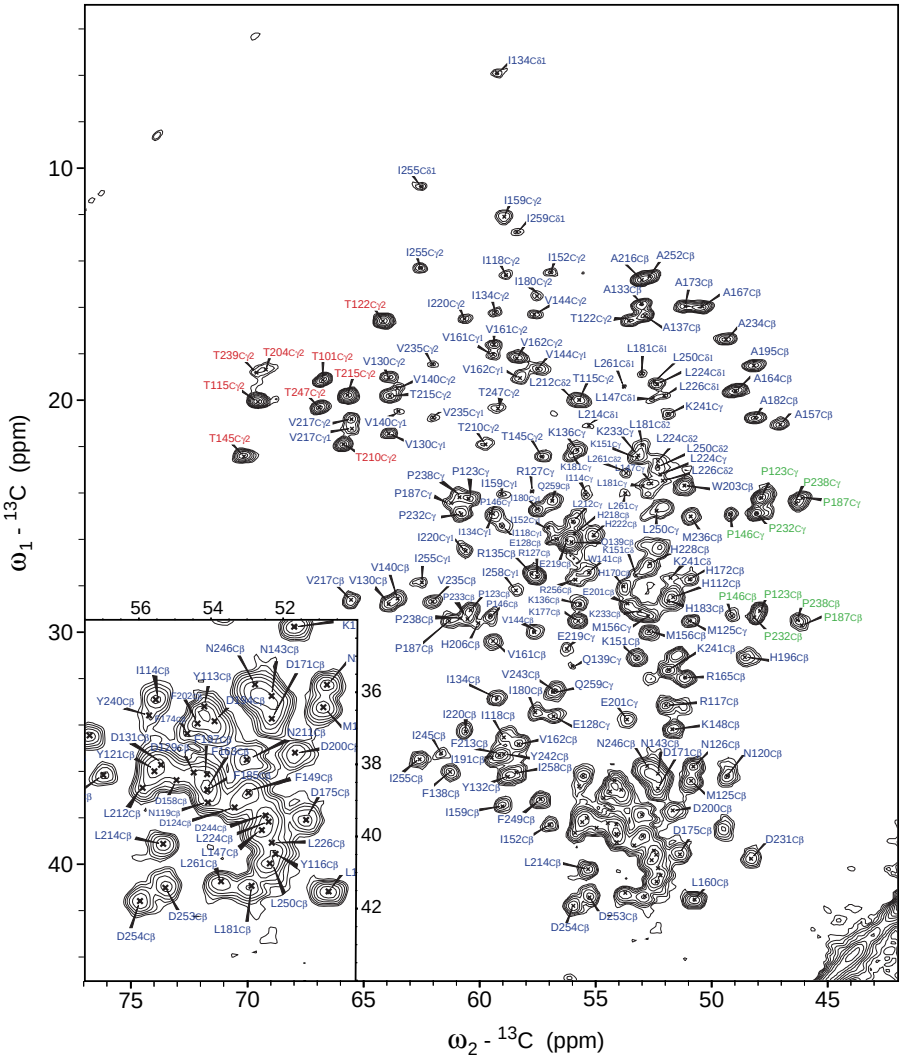


Figure 4. Particular of the aliphatic region of the ^{13}C–^{13}C PDSD spectrum with reported assignment. The blue, red, and green colors indicate the correlation among sidechain resonances and Cα (blue), Cβ (red) and Cδ (green) resonances.

Table 1. Percentage assignment of the microcrystalline Zn-MMP-12.[a]

	C^{α}	$C^{sidechain}$	$C^{aliphatic}$	C'	total
PDSD	76 %	75 %	75 %	48 %	68 %
+ J-dec[b]	88 %	83 %	86 %	70 %	81 %
+ 3D[c]	96 %	90 %	92 %	93 %	93 %

[a] As the assignments of the terminal N- and C- part of the protein are not available in solution, all the percentages are referred to the 151 AA for which the solution assignment is available.

[b] ^{13}C–^{13}C PDSD J-decoupled.

[c] 3D NCACX and 3D NCOCX.

Acknowledgements

This work was financially supported by the EU contracts: ORTHO AND PARA WATER n°005032, EU-NMR Contract n° 026145 (JRA 2 HET-NMR), and by the project MIUR - RBAU013NSB, and by Ente Cassa di Risparmio di Firenze.

Keywords: NMR spectroscopy · solid-state NMR · protein structure · protein NMR · MMP-12

[1] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, H. Oschkinat, *Nature* **2002,** *420,* 98-102.

[2] A. Lange, K. Giller, S. Hornig, M. F. Martin-Eauclaire, O. Pongs, S. Becker, M. Baldus, *Nature* **2006,** *440,* 959-962.

[3] T. I. Igumenova, A. E. McDermott, K. W. Zilm, R. W. Martin, E. K. Paulson, A. J. Wand, *J.Am.Chem.Soc.* **2004,** *126,* 6720-6727.

[4] F. Castellani, B. J. van Rossum, A. Diehl, K. Rehbein, H. Oschkinat, *Biochemistry* **2003,** *42,* 11476-11483.

[5] J. Pauli, B. van Rossum, H. Forster, H. J. de Groot, H. Oschkinat, *J.Magn Reson.* **2000,** *143,* 411-416.

[6] A. E. Bennett, R. G. Griffin, J. H. Ok, S. Vega, *J. Chem. Phys.* **1992,** *96*, 8624-8627.

[7] C. P. Jaroniec, C. E. MacPhee, V. S. Bajaj, M. T. McMahon, C. M. Dobson, R. G. Griffin, *Proc.Natl.Acad.Sci.U.S.A* **2004,** *101,* 711-716.

[8] A. Lange, S. Luca, M. Baldus, *J Am.Chem Soc.* **2002,** *124,* 9704-9705.

[9] A. Bockmann, *Comptes Rendu Chimie* **2006,** *9,* 381-392.

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

[11] S. K. Straus, T. Bremi, R. R. Ernst, *Chem.Phys.Lett.* **1996,** *262,* 709-715.

[12] J. Pauli, M. Baldus, B. van Rossum, H. de Groot, H. Oschkinat, *Chembiochem.* **2001,** *2,* 272-281.

[13] T. I. Igumenova, A. J. Wand, A. E. McDermott, *J.Am.Chem.Soc.* **2004,** *126,* 5323-5331.

Received: ((will be filled in by the editorial staff))
Published online: ((will be filled in by the editorial staff))

((Author(s), Corresponding Author(s)))* ((Insert TOC Graphic here)) ((Text for Table of Contents))

((Title))

Reference List

[1] F. Castellani, B. van Rossum, A. Diehl, M. Schubert, K. Rehbein, H. Oschkinat, *Nature* **2002,** *420* 98-102.

[2] A. Lange, K. Giller, S. Hornig, M. F. Martin-Eauclaire, O. Pongs, S. Becker, M. Baldus, *Nature* **2006,** *440* 959-962.

[3] T. I. Igumenova, A. E. McDermott, K. W. Zilm, R. W. Martin, E. K. Paulson, A. J. Wand, *J.Am.Chem.Soc.* **2004,** *126* 6720-6727.

[4] F. Castellani, B. J. van Rossum, A. Diehl, K. Rehbein, H. Oschkinat, *Biochemistry* **2003,** *42* 11476-11483.

[5] J. Pauli, B. van Rossum, H. Forster, H. J. de Groot, H. Oschkinat, *J.Magn Reson.* **2000,** *143* 411-416.

[6] A. E. Bennett, R. G. Griffin, J. H. Ok, S. Vega, *J Chem Phys* **1992,** *96* 8624-8627.

[7] C. P. Jaroniec, C. E. MacPhee, V. S. Bajaj, M. T. McMahon, C. M. Dobson, R. G. Griffin, *Proc.Natl.Acad.Sci.U.S.A* **2004,** *101* 711-716.

[8] A. Lange, S. Luca, M. Baldus, *J Am.Chem Soc.* **2002,** *124* 9704-9705.

[9] A. Bockmann, *Comptes Rendu Chimie* **2006,** *9* 381-392.

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

[11] S. K. Straus, T. Bremi, R. R. Ernst, *Chem.Phys.Lett.* **1996,** *262* 709-715.

[12] J. Pauli, M. Baldus, B. van Rossum, H. de Groot, H. Oschkinat, *Chembiochem.* **2001,** *2* 272-281.

[13] T. I. Igumenova, A. J. Wand, A. E. McDermott, *J.Am.Chem.Soc.* **2004,** *126* 5323-5331.

Solid-State NMR of Matrix Metalloproteinase 12: an Approach Complementary with Solution NMR

Stéphane Balayssac,[a] Ivano Bertini,*[a],[b] Katja Fälber,[c] Marco Fragai,[a],[d] Stefan Jehle,[c] Moreno Lelli,[a] Claudio Luchinat,[a],[d] Hartmut Oschkinat,[c] and Kwon Joo Yeo[a],[b]

[a] Magnetic Resonance Center (CERM), University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino (Italy)
[b] Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino (Italy)
[c] Leibnizinstitut für Molekulare Pharmakologie, Robert-Rössle-Straße 10, 13125 Berlin (Germany)
[d] Department of Agricultural Biotechnology, University of Florence, Via Maragliano 75-77, 50144 Florence (Italy).

Supporting Information

Materials and Methods

Solid-State NMR spectroscopy: All the NMR spectra were recorded on a Bruker Avance 700 wide bore instrument operating at 16.4 T (700 MHz of ^{1}H Larmor frequency and 176.0 MHz of ^{13}C Larmor frequency). A double-channel and triple-channel 4 mm CP-MAS probehead was used. The spinning frequency of the 4mm ZrO_2 HR MAS rotors was stabilized to ±2 Hz. Solid-state chemical shifts are referred to DSS following the procedure reported in literature.[1] The probe temperature was kept to 280 K, which ensure a sample temperature around 290 K.

In double resonance CP experiments the ^{1}H 90° pulse was set to 3.5 μs, during cross-polarization the ^{1}H B_1 was 61 kHz with a mixing time of 0.75 ms. A 100%/50% ramp was applied on the carbon channel with a 100% power level of 46 kHz.

Similar parameters were used for a standard proton-driven spin diffusion sequence (2D PDSD); the ^{13}C 90° pulse was 4.2 μs, and the mixing times were varied from 5 to 15 ms. The MAS frequency was varied from 8.5 to 11.5 kHz. The SPINAL-64[2] decoupling at 71 kHz of power was used on proton, switching off the decoupling during the mixing time. All the experiments were acquired with 16 scans per experiment, with 1536 direct points, and 1280 experiments in the State-TPPI scheme.

The same parameters and MAS frequency were used for the J-decoupled PDSD, using a 350 μs gauss shaped pulse for the selective Cα inversion. This spectrum was acquired with 48 scans, 2048 direct points and 384 experiments using the States-TPPI scheme.

The triple resonance 3D NCACX PDSD and 3D NCOCX PDSD experiments were acquired using the standard sequences.[3] The ^{13}C 90° pulse was 5.2 μs. The HN Hartmann-Hahn matching was optimized with ^{1}H 90° of 3.5 μs, 37 kHz for the ^{15}N B_1 and a 100%/50% ramp on the ^{1}H channel

applied for 1.8 ms at 60 kHz of 100% power. The NC matching was optimized using the ^{15}N B_1 of 35 kHz and a Tangent Amplitude Modulated ramp at 25 kHz on ^{13}C. The NC matching was 6.25 ms long using a simultaneous ^{1}H CW decoupling of 94 kHz. SPINAL64 ^{1}H decoupling at 71 kHz was used during the direct and indirect acquisition times. The A weak ^{1}H CW radio frequency optimized at 10 kHz was used during the 35 ms of PDSD mixing time. Both NCACX and NCOCX experiments were recorded at 11 kHz MAS frequency, with 64 scans per experiment, 1536 direct points, and 96 increments in the indirect ^{13}C dimension and 32 increments in the indirect ^{15}N dimension using the TPPI scheme for both the indirect dimensions.
Parameters similar to that used for the 3D experiments were used also for the 2D NCA spectrum.[3] 2D data were processed using zero-filling up to 4096 points in the direct dimension and 2048 points in indirect dimension, and using gaussian and square cosine filters for direct and indirect dimension, respectively.

Preparation of the microcrystalline Zn-MMP-12 sample: The catalytic domain of Zn-MMP-12 protein with NNGH (N-isobutyl-N-[4-methoxyphenylsulphonyl]glycyl hydroxamic acid) inhibitor was expressed and crystallized following the already published procedure.[4] The microcrystals thus obtained were washed in 920 µl of a low-salt reservoir buffer (10 mM Tris·HCl/25% PEG 8000 pH 8) for 1 day. Approximately 20 mg of these crystals were transferred in a 50 µl 4mm ZrO_2 HR MAS rotor, and sealed with the upper spacer in order to maintain the amount of water constant.

Figures

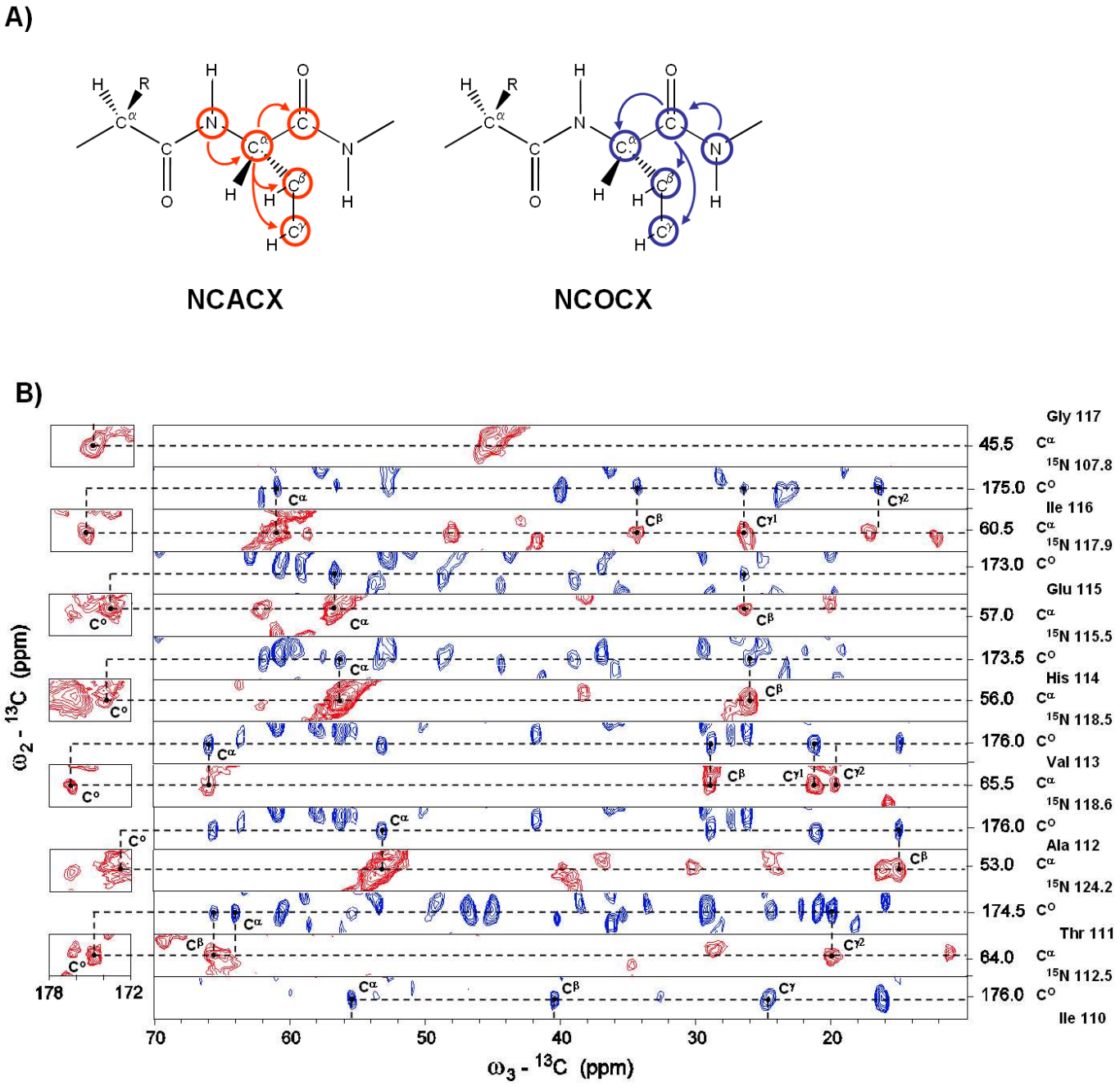


Figure S1. A) Representation of the NCACX (red) and NCOCX (blue) correlation experiments (16.4 T, $\omega_R/2\pi$ = 11 kHz). Le circles and arrows show detected nuclei and the correlation scheme, respectively. B) Sequential specific assignment using NCACX (red) and NCOCX (blue) spectrum for 8 residues (Ile110-Gly117) of the microcrystalline catalytic domain of MMP-12.

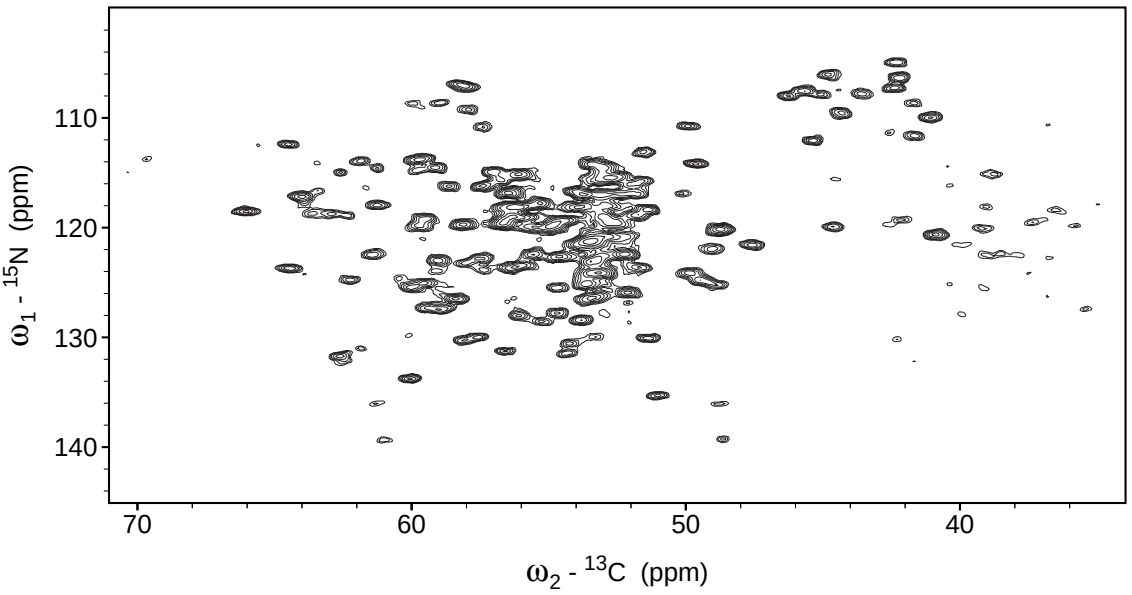


Figure S2. 2D Adiabatic-CP NCA correlation experiment of the Zn-MMP-12 (16,4 T, $\omega_R/2\pi$ = 11 kHz). Weak N-C_β correlation peaks are also observable.

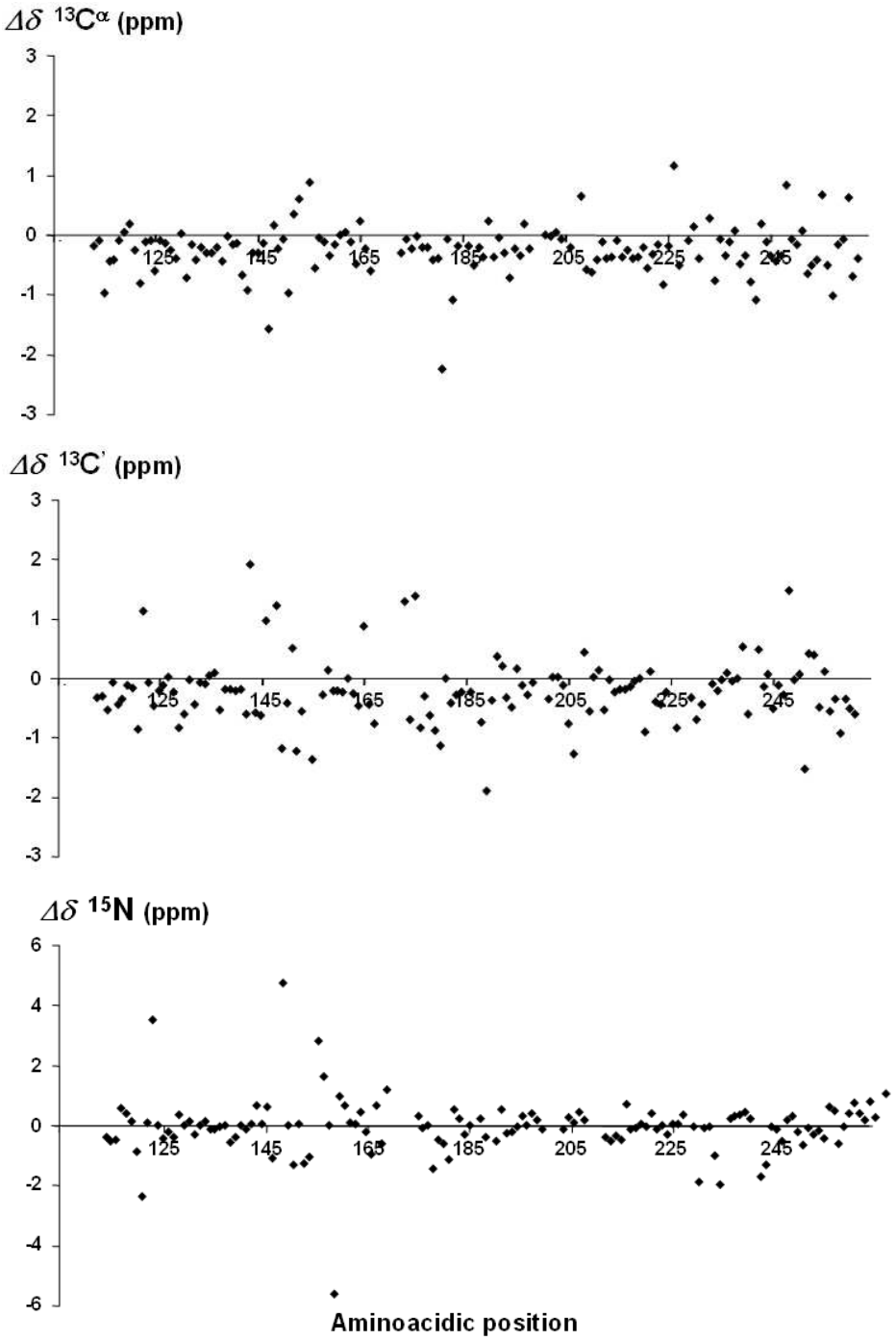


Figure S3. Plot of the $^{12}C^{\alpha}$, $^{13}C'$ and ^{15}N chemical shift differences among solid-state (δ_S) and liquid-state (δ_L) chemical shifts ($\Delta\delta = \delta_S - \delta_L$), for each residue positions. The Root Mean Square Deviation (RMSD) is 0.47, 0.56, and 1.19 ppm for $^{13}C_{\alpha}$, $^{13}C'$ and ^{15}N respectively.

Assignment Tables

Table S1. ^{13}C and ^{15}N chemical shifts values (ppm) of the microcrystalline precipitate of the catalytically active domain of Zn-MMP-12 with NNGH.

Residue	AA	^{15}N	^{13}CO	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C^{\gamma}$	$^{13}C^{\delta}$	$^{13}C^{\varepsilon}$
HIS	112		173.1	54.1	31.0			
TYR	113	119.9	173.6	56.5	38.9			
ILE	114	128.0	174.8	57.9	38,6	26.6/16.9	11.1	
THR	115	116.9	174.5	58.1	72.1	22.5		
TYR	116	118.0	171.3	54.6	43.0			
ARG	117	119.5	173.9	54.3	35.6	27.2	40.3	
ILE	118	127.3	175.0	61.4	37.0	27.9/17.0	13.7	
ASN	119	127.8	173.7	56.5	41.5	174.3		
ASN	120	110.8	169.0	51.7	38.7	177.6		
TYR	121	115.2	175.7	57.9	40.7			
THR	122	119.3	173.4	55.9	66.6	19.1		
PRO	123		176.7	62.8	31.5	26.7	50.3	
ASP	124	121.5	175.7	55.7	41.7	183.1		
MET	125	113.1	174.7	53.3	38.9	32.0		
ASN	126	118.4	177.4	53.2	38.3	176.0		
ARG	127	126.5	177.8	60.1	30.0	26.3	42.6	
GLU	128	116.2	178.6	59.2	28.5	36.1	176.4	
ASP	129	120.0	178.6	56.9	40.7			
VAL	130	123.7	177.2	66.3	31.2	23.9/21.5		
ASP	131	118.4	179.2	57.8	40.5			
TYR	132	119.8	176.3	61.2	38.6			
ALA	133	121.2	179.3	55.3	18.4			
ILE	134	113.8	177.3	61.7	35.3	27.4/18.7	8.4	
ARG	135	119.7	179.9	59.9	30.1	27.4	43.6	
LYS	136	119.0	178.3	58.1	31.2	24.5	27.9	
ALA	137	123.0	178.5	55.2	18.8			
PHE	138	113.9	178.6	63.6	38.5			
GLN	139	119.1	177.7	58.3	28.6	34.0		
VAL	140	117.2	178.2	65.9	31.0	23.0/21.9		
TRP	141	116.9	180.3	57.8	29.9			
SER	142	119.2	177.5	60.5	63.2			
ASN	143	115.4	176.2	54.7	38.5	172.5		
VAL	144	107.1	173.9	60.0	32.4	21.1/18.8		
THR	145	109.3	174.1	59.7	72.6	24.9		
PRO	146	129.8		61.9	31.8	27.4	51.5	
LEU	147	115.4	177.6	55.0	42.4	26.2	22.5/24.5	
LYS	148	122.4	172.6	54.0	36.6	25.3	30.2	41.5
PHE	149	118.1	176.2	55.4	41.2			
SER	150	116.7	171.5	55.7	65.1			
LYS	151	126.4	175.4	55.6	33.6	25.0	30.0	42.2
ILE	152	122.8	174.9	59.4	40.8	27.9/17.0	13.9	
ASN	153							
THR	154	114.5	171.3	61.1	71.2	21.8		
GLY	155	111.6	173.7	43.6				
MET	156	119.1	174.7	55.1	32.4	31.8		
ALA	157	121.5	176.5	49.4	23.5			
ASP	158	122.4	176.8	57.4	40.9	179.4		
ILE	159	125.0	173.1	61.4	39.9	26.6/14.7	14.6	
LEU	160	130.1	175.5	53.2	44.0	27.2	23.9/25.3	
VAL	161	125.5	175.1	61.8	32.9	20.6/20.1		
VAL	162	127.5	174.5	60.7	37.2	21.4/20.7		
PHE	163	125.5	175.7	56.4	41.2			
ALA	164	125.0	175.7	51.3	22.1			

ARG	165	118.8	175.4	53.5	34.4	26.3	43.4
GLY	166	107.5	175.1	45.9			
ALA	167	135.3	176.3	52.8	18.5		
HIS	168						
GLY	169		170.1	44.4			
ASP	170						
ASP	171	120.8	175.5	54.6	39.2	179.9	
HIS	172		174.0	53.3	30.2		
ALA	173	123.6	179.6	53.5	18.4		
PHE	174	119.9	177.6	57.1	39.6		
ASP	175	116.4	176.6	53.6	42.0	179.8	
GLY	176	108.6	172.9	43.5			
LYS	177	123.6	177.2	58.2	32.0	24.6	29.3
GLY	178	120.6	173.3	42.5			
GLY	179	109.6	173.6	46.0			
ILE	180	130.3	177.0	60.0	35.9	27.2/18.0	10.0
LEU	181	128.4	175.6	55.4	43.9	26.2	21.4/24.4
ALA	182	114.2	175.6	50.7	23.2		
HIS	183	116.7	171.6	54.1	31.1		
ALA	184						
PHE	185	131.4	175.4	56.5	41.1		
GLY	186	105.0	175.4	43.9			
PRO	187	131.1	174.0	63.7	32.1		48.6
GLY	188	106.4	171.5	43.8			
SER	189	115.0	175.7	58.7	64.3		
GLY	190	112.1	175.4	46.8			
ILE	191	133.8	174.6	61.8	37.9	27.9/16.3	14.2
GLY	192	106.0	174.1	46.4			
GLY	193	119.9	171.6	46.2			
ASP	194	122.6	173.7	56.2	39.4	179.9	
ALA	195	120.2	174.5	50.5	21.0		
HIS	196	122.0	172.5	50.8	33.6		
PHE	197	122.7	174.0	56.5	40.8		
ASP	198						
GLU	199						
ASP	200		179.2	54.1	40.1		
GLU	201	130.6	176.6	56.0	31.4	36.2	183.3
PHE	202	119.6	177.3	56.7	39.3		
TRP	203	116.9	174.7	53.6	26.2		
THR	204	126.5	174.7	59.6	71.8	21.1	
THR	205	123.0	173.7	60.7	69.3	21.6	
HIS	206		175.5	62.1	32.0		
SER	207		176.4	58.9	63.1		
GLY	208	107.3	173.0	44.4			
GLY	209	107.8	173.7	45.2			
THR	210	124.6	172.7	62.1	68.3	24.3	
ASN	211	125.1	175.9	55.3	40.4	180.7	
LEU	212	131.3	176.3	58.3	41.2	27.7	27.4/22.4
PHE	213	121.0	175.3	61.4	37.8		
LEU	214	118.2	178.7	57.8	42.7	26.8	23.6/26.6
THR	215	112.4	177.0	66.3	68.2	22.3	
ALA	216	124.2	178.5	55.4	17.3		
VAL	217	118.5	178.7	67.9	31.1	23.8/23.2	
HIS	218	118.5	176.1	58.4	28.4		
GLU	219	115.6	176.8	58.6	28.6	33.3	176.2
ILE	220	118.0	177.4	63.0	36.7	29.0/18.9	12.0
GLY	221	107.9	176.9	47.9			
HIS	222	119.7	180.4	57.5	28.3		
SER	223	122.5	175.6	55.8	63.0		
LEU	224	114.6		54.8	42.2	25.8	21.8/25.1
GLY	225		173.4	45.5			
LEU	226	120.9		54.6	42.7	22.3	22.3/26.9

GLY	227	107.9	174.9	46.8				
HIS	228	115.5	174.9	55.1	29.6			
SER	229	114.2	174.8	55.2	65.1			
SER	230	118.9	173.0	58.2	64.0			
ASP	231	124.8	174.9	50.8	42.2	180.2		
PRO	232		177.7	63.1	31.9	27.4	50.5	
LYS	233	116.9	176.6	55.5	31.7	24.8	28.2	
ALA	234	124.1	179.7	51.7	19.8			
VAL	235	131.8	179.1	64.4	31.1	23.2/20.9		
MET	236	115.9	177.2	53.3	27.5	30.0		
PHE	237	128.6		56.6	39.1			
PRO	238		174.7	63.2	31.9	26.8	49.1	
THR	239	114.7	172.7	61.6	72.0	21.9		
TYR	240	126.6		58.1	39.1			
LYS	241	129.9	173.9	54.4	34.3	23.0	30.1	42.0
TYR	242	124.0	174.8	61.1	37.8			
VAL	243	123.8	172.2	59.1	34.9	21.3/19.5		
ASP	244	119.2	178.6	54.7	41.9			
ILE	245	124.7	177.1	64.0	37.6	26.8/17.7	13.4	
ASN	246	118.2	176.2	55.1	38.3	173.5		
THR	247	108.7	175.3	61.7	69.4	22.6		
PHE	248	123.2	174.4	59.7	39.7			
ARG	249	125.9	174.8	53.9	33.6	26.0	43.3	
LEU	250	120.9	175.6	54.8	43.2	27.2	21.7/25.4	
SER	251	117.8	175.4	57.0	65.6			
ALA	252	124.2	180.5	55.0	17.2			
ASP	253	118.1	178.2	57.7	43.9	179.7		
ASP	254	118.2	177.9	58.3	44.2	180.6		
ILE	255	118.8	177.1	65.0	37.9	30.3/16.8	13.2	
ARG	256	119.5	180.2	58.3	30.2	27.1	43.2	
GLY	257	107.6	176.2	47.2				
ILE	258	124.0	177.6	60.8	38.6	30.6/20.5	15.2	
GLN	259	123.7	178.8	59.2	26.8	35.0	181.7	
SER	260	116.3	174.3	60.6	63.0			
LEU	261	119.4		56.1	43.8	26.4	21.9/25.6	
TYR	262							
GLY	263	111.2	171.2	44.0				

Table S2. ^{13}C and ^{15}N chemical shifts values (ppm) of Zn-MMP-12 catalytically active domain with NNGH in solution.

Residue	AA	^{15}N	^{13}CO	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	$^{13}C^{\gamma}$	$^{13}C^{\delta}$	$^{13}C^{\varepsilon}$
HIS	112		173.4	54.3	31.5			
TYR	113	120.3	173.9	56.6	38.2			
ILE	114	128.6	175.3	58.8	40.2	27.8/17.0	12.9	
THR	115	117.4	174.6	58.5	72.6	22.0		
TYR	116	117.4	171.7	55.0	43.2			
ARG	117	119.2	174.2	54.4	34.9	28.9	42.1	
ILE	118	127.3	175.1	61.4	36.5	27.7/16.5	12.9	
ASN	119	128.7	173.8	56.3	41.7	175.3		
ASN	120	113.2	169.8	51.9	38.5	177.8		
TYR	121	115.1	174.6	58.7	40.7			
THR	122	115.8	173.5	56.0	67.0			
PRO	123		177.1	62.9	31.8		50.3	
ASP	124	121.9	175.9	56.3	41.8	183.2		
MET	125	113.4	174.8	53.4	39.1	32.0		
ASN	126	118.8	177.4	53.3	38.3	176.0		
ARG	127	126.2	178.0	60.4	29.6	26.3	43.3	
GLU	128	116.2	179.4	59.6	28.5	36.1		
ASP	129	119.9	179.2	56.9	40.7			
VAL	130	124.1	177.2	67.0	31.2	23.8/21.1		
ASP	131	118.4	179.6	58.0	40.4			
TYR	132	119.7	176.4	61.6	38.7			
ALA	133	121.4	179.4	55.5	18.5			
ILE	134	114.0	177.2	62.0	35.3	27.3/18.0	8.2	
ARG	135	119.8	179.8	60.2	30.0	27.2	44.0	
LYS	136	119.0	178.8	58.3	31.0	24.1	27.8	41.8
ALA	137	123.6	178.7	55.6	19.0			
PHE	138	114.4	178.8	63.6	38.7			
GLN	139	119.1	177.9	58.5	28.4	34.2		
VAL	140	117.3	178.4	66.0	31.1	23.0/21.1		
TRP	141	116.9	180.9	58.4	29.8			
SER	142	118.6	175.7	61.4	63.5			
ASN	143	115.3	176.7	55.0	38.8			
VAL	144	106.5	174.5	60.3	32.7	21.5/19.3		
THR	145	110.4	173.2	59.9	72.2	24.5		
PRO	146		175.5	63.4	32.3			
LEU	147	110.7	176.4	54.9	42.9			
LYS	148	122.4	173.7	54.2	35.0	24.4	28.5	42.2
PHE	149	119.5	176.6	55.5	41.7			
SER	150	116.7	171.0	56.6	65.7			
LYS	151	127.7	176.6	55.3	33.3	24.8	29.6	42.0
ILE	152	123.9	175.4	58.8	40.5	26.0/17.3	12.8	
ASN	153		175.0	53.6	39.9			
THR	154	111.8	172.6	60.3	70.9	21.5		
GLY	155	110.0		44.1				
MET	156	119.1	175.0	55.1	31.8			
ALA	157	127.2	176.4	49.5	22.0			
ASP	158	121.5	177.0	57.8	41.5	179.4		
ILE	159	124.4	173.3	61.6	40.2	26.0/14.7	14.5	
LEU	160	130.1	175.7	53.2	43.5		24.4/26.2	
VAL	161	125.4	175.1	61.8	32.8	20.4/20.0		
VAL	162	127.1	174.8	60.8	37.2	21.1/20.8		
PHE	163	125.7	176.1	56.9	40.8			
ALA	164	126.0	174.9	51.1	22.6			
ARG	165	118.2	175.8	53.7	33.6	26.1	43.1	
GLY	166	108.1	175.8	46.5				

ALA	167	134.1		53.1	17.9			
HIS	168		174.1					
GLY	169							
ASP	170		174.1	56.9	40.7			
ASP	171	121.6						
HIS	172	120.6	172.8	53.6	30.0			
ALA	173	123.3	180.3	53.6	18.3			
PHE	174	120.0	176.3	57.3	38.7			
ASP	175	116.4	177.4	53.6	41.6	179.6		
GLY	176	110.1	173.2	43.7				
LYS	177	124.1	177.8	58.4	32.0	24.9	28.8	42.2
GLY	178	121.2	174.1	42.9				
GLY	179	110.7	174.7	46.4				
ILE	180	129.8	177.0	62.1	38.0	28.1/18.0	11.9	
LEU	181	128.2	176.0	55.5	43.9	26.0	21.0/24.4	
ALA	182	114.5	175.9	51.7	22.1			
HIS	183	116.7	171.8	54.3	31.0			
ALA	184	119.9	179.6	53.6	18.1			
PHE	185	131.2	175.6	56.7	41.3			
GLY	186	105.4		44.4				
PRO	187		174.7	63.9	31.8	34.2		
GLY	188	106.9	173.3	44.2				
SER	189	114.5	176.1	58.5	64.2			
GLY	190	112.4	175.0	47.2				
ILE	191	134.0	174.4	61.9	38.0	27.7/16.0	13.9	
GLY	192	106.1	174.4	46.7				
GLY	193	119.6	172.1	46.9				
ASP	194	122.6	173.6	56.4	39.7	180.1		
ALA	195	119.8	174.6	50.8	21.1			
HIS	196	121.8	172.8	50.6	33.7			
PHE	197	122.8	174.1	56.7	40.9			
ASP	198		176.0	50.6	23.2			
GLU	199		176.4	59.3	35.5			
ASP	200		179.5	54.1	40.2			
GLU	201	130.7	176.6	56.0	31.0	36.4	183.2	
PHE	202	119.4	177.3	56.7	39.3			
TRP	203	116.8	174.8	53.7	26.1			
THR	204	126.0	175.4		72.5			
THR	205	122.8	174.9	60.9	69.3	22.4		
HIS	206	120.7						
SER	207	123.6	176.0	58.3	63.4			
GLY	208	115.3	173.5	45.0				
GLY	209	108.2	173.7	45.8				
THR	210	125.2	172.6	62.5	68.5	23.8		
ASN	211	125.5	176.4	55.4	41.7	180.2		
LEU	212	131.8	176.3	58.7	41.1	27.7	27.7/22.3	
PHE	213	120.3	175.5	61.8	37.7			
LEU	214	118.3	178.9	57.9	43.1	26.2	23.8/25.9	
THR	215	112.6	177.2	66.7	68.3	21.7		
ALA	216	124.2	178.6	55.7	17.3			
VAL	217	118.6	178.7	68.3	31.1	23.5/23.3		
HIS	218	118.2	176.1	58.8	28.6			
GLU	219	115.7	177.7	58.8	28.6	33.3		
ILE	220	118.0	177.3	63.5	37.2	29.1/18.8	11.9	
GLY	221	108.2	177.3	48.2				
HIS	222	119.7	180.8	57.7	28.3			
SER	223	122.5	175.8	56.6	63.0			
LEU	224	114.3	178.0	55.0	42.2	24.2	22.1/24.6	
GLY	225		174.2	44.4				
LEU	226	121.0	177.5	55.1	42.5	21.8	26.3/21.8	
GLY	227	109.8						
HIS	228	115.7	175.2	55.2	29.2			

SER	229	114.3	175.4	55.1	65.3			
SER	230	119.9	173.4	58.6	64.0			
ASP	231	126.8			42.0			
PRO	232		177.8	62.9	31.9		50.6	
LYS	233	116.7	176.8	56.2	31.9	24.9	28.2	
ALA	234	123.9	179.7	51.8	19.6			
VAL	235	131.4	179.0	64.8	31.1	23.0/20.7		
MET	236	115.5	177.2	53.4	27.0			
PHE	237	128.4		56.5	39.9			
PRO	238		174.2	63.7				
THR	239	116.5	173.3	61.9	71.8	21.9		
TYR	240	128.0	174.1	58.9	38.9			
LYS	241	129.9	173.4	55.4	34.5	24.0	28.8	42.1
TYR	242	124.1	174.9	60.9	38.3			
VAL	243	124.4	172.1	59.2	35.1	21.1/19.6		
ASP	244	119.1	179.1	55.0	41.9			
ILE	245	124.4	177.2	64.4	37.6	27.0/17.8	13.5	
ASN	246	118.5	176.5	55.4	38.3			
THR	247	109.4	173.9	60.9	69.4	21.1		
PHE	248	123.3	174.4	59.8	39.9			
ARG	249	126.2	174.7	54.1	33.4	26.0	43.4	
LEU	250	121.1	177.0	54.7	42.9	26.7	22.4/26.5	
SER	251	118.3	175.0	57.6	65.9			
ALA	252	123.6	180.1	55.5	17.8			
ASP	253	117.7	178.7	58.1	43.9	180.2		
ASP	254	118.8	177.8	57.7	44.2	180.0		
ILE	255	118.9	177.6	65.5	38.2	30.5/16.6	12.5	
ARG	256	119.2	180.5	59.3	29.8	27.1	43.2	
GLY	257	106.8	177.1	47.4				
ILE	258	123.6	177.9	60.9	38.5	30.7/20.6	14.9	
GLN	259	123.5	179.3	58.6	27.2	34.9	180.7	
SER	260	115.5	174.9	61.3	62.9			
LEU	261	119.1	176.7	56.5	43.5	25.4	21.5/25.6	
TYR	262		176.0	59.8				
GLY	263	110.2						

[1] C.R. Morcombe, K.W. Zilm, *J. Magn. Reson.* **2003,** *162*, 479−486.

[2] B.M. Fung, A.K. Khitrin, K. Ermolaev, *J. Magn. Reson.* **2000,** *142*, 97−101.

[3] F. Castellani, B.J. van Rossum, A. Diehl, K. Rehbein, H. Oschkinat, *Biochemistry* **2003**, *42*, 11476−11483.

[4] I. Bertini, V. Calderone, M. Cosenza, M. Fragai, Y.M. Lee, C. Luchinat, S. Mangani, B. Terni, P. Turano, *Proc. Natl. Acad. Sci. U S A.* **2005,** *102*, 5334−5339.

5 General discussion and perspectives

5.1 General discussion

The catalytic domains of MMP-1, MMP-7, MMP-8, MMP-12, the full length MMP-12 (Glu219Ala), and the hemopexin domain of MMP-12 were prepared from *Escherichia coli* strains. The proteins have been chosen according to a general strategy aimed to select some of the most pathologically relevant matrix metalloproteinases. Thanks to the methods and strategies developed during this PhD project a large amount of concentrated protein samples have been produced. These samples allowed us to deeply investigate the structural features of the selected catalytic domains of matrix metalloproteinases and to carefully analyze the interaction of these proteins with inhibitors and substrate models.

Concerning the fibroblast collagenase, its expression is detected in a variety of physiological processes including embryonic development, wound healing, as well as in a number of pathological processes, including arthritis, cardiovascular disease and different types of malignant tumors. Therefore, several candidate drugs have been designed and successfully tested in vitro. Unfortunately, they failed the clinical trials due to a lack of selectivity. The structural analysis of several protein-ligand adducts would be of great help in order to increase the selectivity towards this important target. During this project two different protein constructs (Asn106-Gly261 and Pro21-Pro269) have been designed and expressed in order to deeply investigate interactions of the protein with ligands. Although all attempts to develop a reproducible crystallization protocol failed, nevertheless the methodological advancements allowed us to produce large amounts of stable and soluble MMP-1 samples which have been employed to design a new approach for protein-ligand structural analysis. In this approach a combination of *in silico* tools and experimental NMR data was proposed for a relatively fast determination of protein-ligand structural models (Chapter 4.1). The strategy that we proposed promises to be useful also for the structural

determination of different protein-ligand adducts, whenever the structure of the free protein is known and the structural changes upon complexation are not expected to be dramatic.

In a different research line, the catalytic domain of MMP-1 (activated from pro-cat MMP-1) has been also exploited to develop an high-throughput screening protocol based on SPR techniques. It is currently believed that MMPs can not be anchored on the sensor-chip surface without losing their activity. Therefore our efforts have been devoted to develop a strategy to monitor the activity of the anchored protein. Several samples of MMP-1 have been used to develop the experimental protocol that coupling the Fourier Transform-SPR (FT-SPR) technique with Electrospray Ionization-Mass Spectroscopy (ESI-MS) allows us to evaluate the enzymatic activity of MMP-1 catalytic domain anchored on gold surfaces (Chapter 4.2). The results we got demonstrate the possibility of monitoring enzyme activity and/or interactions with other molecules. The same approach has been extended to MMP-12 and MMP-8. The strategy has been further implemented by using the *in situ* Atmospheric Pressure MALDI-MS (Chapter 4.3). We suppose that the new experimental procedure is particularly suitable for SPR/MALDI-MS coupled investigation of enzyme activity and/or to monitor the interactions with other molecules.

Also the macrophage metalloelastase (MMP-12) is a relevant target for drug design. Under pathological conditions MMP-12 is involved in emphysema and multiple sclerosis. The enzyme is also able to degrade ECM components such as elastin and collagen type IV and involved tissue remodeling. To accomplish rational drug design for the protein, high resolution structures of the protein with ligands are essential. For this purpose, during this PhD study we developed methods to obtain large amounts of the catalytic domain of MMP-12 from *Escherichia coli.* These samples have been employed for NMR studies and for crystals preparation. In particular the availability of the stable samples made possible to collect structural data on adducts of the collagen fragment ProGlnGlyIleAlaGly peptide with the protein and on a large number of adducts of a homologous series of ligands with the

catalytic domain of MMP-12. The reaction mechanism of the catalytic domain of MMPs was well provided by crystal structures of the uninhibited form of the protein and of the adducts with the collagen fragment peptide (Chapter 4.4).

Thanks to the soaking protocols developed during the project the structural details of the adducts of MMP-12 with 14 structural-related inhibitor have been deeply analyzed. From combining the structural information with calorimetric and inhibition measurements we obtained hints for the planning of new finely tuned inhibitors for MMP-12 (Chapter 4.5).

The expertise acquired on MMPs has been of fundamental importance also to speed-up our advancements in solid-state NMR. The recent results in solid-state NMR demonstrate the potential of this technique for solving membrane protein structure, one of the major challenges in structural biology today, or for getting information on protein aggregates. Two of the main limitations in solid-state NMR are related to the large amount of ^{13}C- and ^{15}N-labeled sample needed for the measurement and to the physical state of the sample (only microcrystalline samples ensure high quality spectra). The high yield of the MMP-12 and the quality of the microcrystals allowed us to obtain well-resolved 2D and 3D spectra that have been used for assignment of the protein. In the attached article (Chapter 4.6) we demonstrate that it is possible to translate in a short time a large fraction of the ^{13}C liquid assignment to solid-state. This opens the possibility to investigate proteins involved in important biological processes such as oligomerization or protein-protein interactions, without the line broadening associated to the increase in molecular weight present in solution.

Although the studies performed on the catalytic domains can provide important information on the dynamical properties, on the enzymatic mechanism and on the structural features of this important class of proteins, only the investigation on the active full length proteins provides a full picture of the protein structure and function. At this regards our efforts have been devoted to obtain concentrated sample of MMP-12 full length with a well-

folded state. The new protocol, developed during this PhD project, allowed us to overcome all the problems previously found by the researchers. The difficulty associated to the expression of full length MMPs is demonstrated by the absence in literature of NMR investigation, although two crystal structures have been published. A complete NMR investigation has been already carried out on the active full length MMP-12 (Glu219Ala) and its solution structure is in progress.

Neutrophil collagenase (MMP-8) is another important target belonging to the MMP family. With other collagenases (MMP-1, MMP-13, MT1-MMP (MMP-14), and MMP-2) MMP-8 (collagenase-2, neutrophil collagenase) it is able to degrade type I collagen and it is involved in arthritis and cancer. Therefore also MMP-8 is a potential pharmaceutical target and the investigation of its enzymatic activity can provide valuable data on the degradative mechanism of collagen. Also for the catalytic domain MMP-8 methods to obtain a large mount of the protein for NMR samples and crystals preparation have been developed. Although quality of the crystals can not be compared with that of MMP-12 nevertheless the structures of the free-form of MMP-8 and that obtained in presence of the collagen fragment ProGlnGlyIleAlaGly were of fundamental importance to clarify the reaction mechanism of MMPs (Chapter 4.4). The solution structure of the catalytic domain of MMP-8 in complex with the nanomolar inhibitor NNGH is in progress.

Also MMP-7, matrilysin-1 which lacks the hemopexin domain, has an important role in cancer development. The limited structural information available in literature seems to be related to the difficulties found in protein expression, stability and solubility. Therefore also this protein has been investigated in order to develop new protocols for sample preparation. Although a well-folded protein is currently obtained the solubility remains low so preventing any structural investigation.

5.2 Perspectives

The methods of expression and refolding of several members of human matrix metalloproteinases, developed during this project, will allow the scientific community to produce large amounts of concentrated samples of the catalytic domain of four different MMPs in order to further expand the structural investigation and the biophysical characterization of this family of proteins. In fact to achieve candidate drugs with high affinity and selectivity for each MMP the subtle structural differences among each member of the protein family have to be deeply investigated and elucidated.

Besides these important results and perspectives involving the studies on catalytic domains more exciting expectations are waited from the research line focused on the MMP-12 full length protein. Actually, all domains of active MMP participate to the degradation mechanism of ECM components. Although some theoretical studies on the role of the different domains on the degradation process have been published, only few experimental studies exist in literatures and the mechanism is still unclear. This lack of information is mainly due to the difficulties in “handling” the full length form of these proteins and in preparing concentrated samples. In particular the dramatic self-hydrolysis that affects the full length MMPs is the main limiting factor for the biophysical of studies. The results obtained in the last year have provided the tools to fill this gap so opening new perspectives in MMP research.

In particular the advancements realized in protein engineering, expression and purification allowed us to produce concentrated samples for NMR, x-ray crystallography and biological studies. The structural analysis of the inhibited full length of MMP-12 is being in progress and the results will be available at the beginning of the next year. Moreover new studies with substrate models of elastin and collagen have been already planned and will be carried out during the next year. Obviously this will be the matter of a new PhD project aimed to describe all the steps associated to elastin and collagen

degradation by the full length of MMP-12. Besides the elucidation of the dynamical processes associated to substrate degradation these studies will be relevant for drug design. Actually most of the inhibitors of MMP, designed to target the active site of the catalytic domain, are affected by the lack of selectivity. Therefore the investigation of the whole active protein would provide a new point of view allowing to identify new and protein-specific binding sites and to the design of more selective inhibitors.

6 References

1. Boudreau N. and Bissell M. J., *Curr. Opin. Cell Biol.*, 1998, 10, 640-646
2. Borg T. K., *Am. J. Pathol.* 2004, 164, 1141-1142
3. Werb Z., *Cell*, 1997, 91, 439-442
4. Borkakoti N., *J. Mol. Med*, 2000, 78, 261-268
5. Egeblad M. and Werb Z., *Nat. Rev. Cancer*, 2002, 2, 161-174
6. Coussens L. M., Tinkle C. L., Hanahan D., Werb Z., *Cell*, 2000, 103, 481-490
7. Stetler-Stevenson W. G., *J. Clin. Invest.*, 1999, 103, 1237-1241
8. Coussens L. M., Fingleton B and Matrisian L. M., *Science*, 2002, 295, 2387-2392
9. Lo´pez-Pedrera C., Barbarroja N., Dorado G., Siendones E. and Velasco F., *Leukemia*, 2006, 20, 1331–1340
10. Pei D., *Cancer Cell*, 2005, 7, 207-208
11. Overall C. M. and Kleifeld O., *Nat. Rev. Cancer*, 2006, 6, 227-239
12. Müller-Ladner U., Pap T., Gay R. E., Neidhart M. and Gay S., *Nat. Clin. Pract. Rheumatol*, 2005, 1, 102-110
13. Newby A. C., *Physiol Rev*, 2005, 85, 1–31
14. Li Y. Y., McTiernan C. F., Feldman. A. M., *Cardiovasc. Res.*, 46, 2000, 214–224
15. Nagase H. and Woessner J. F. Jr., *J. Biol. Chem*. 1999, 274, 21491-21494
16. McCawley L. J. and Matrisian L. M., *Curr. Opin. Cell Biol.*, 2001, 13, 534-540
17. Sternlicht M. D. and Werb Z., *Annu. Rev. Cell Dev. Biol*., 2001, 17, 463–516
18. Whittaker M., Floyd C. D., Brown P. and Gearing A. J. H., *Chem. Rev*., 1999, 99, 2735-2776
19. Nagase H., Visse R. and Murphy G., *Cardiovasc. Res.*, 2006, 69, 562-573

20. Andreini, C., Banci, L., Bertini, I., Luchinat, C., and Rosato, A. *J. Proteome. Res.*, 2004, *3*, 21-31

21. Visse, R. and Nagase, H., *Circ. Res.*, 2003, 92, 827-839

22. Massova I., Kotra L. P., Fridman R. and Mobashery S., *FASEB J.* 1998, 12, 1075–1095

23. Remacle A. G., Rozanov D. V., Fugere M., Day R. and Strongin A. Y., *Oncogene*, 2006, 25, 5648-55

24. Lijnen H. R., *Fibrinolysis & Proteolysis*, 2000, 14, 175–181

25. Springman E. B., Angleton E. L., Birkedal-Hansen H. and Van Wart H. E., *Proc. Natl. Acad. Sci. USA*, 1990, 87, 364-368

26. Morgunova E., Tuuttila A., Bergmann U., Tryggvason K., *Proc. Natl. Acad. Sci. U. S. A.*, 2002, 99, 7414-7419

27. Jozic D., Bourenkov G., Lim N. H., Visse R., Nagase H., Bode W., Maskos K., *J. Biol.Chem. 2005*, 280, 9578-9585

28. Ala-aho R., Kähäri V. M., *Biochimie*, 2005, 87, 273-286

29. Aimes R. T., Quigley J. P., *J. Biol. Chem.*, 1995, 270, 5872-5876

30. Tam E. M., Moore T. R., Butler G. S., Overall C. M., *J. Biol. Chem*. 2004, 279, 43336-43344

31. Chung L., Dinakarpandian D., Yoshida N., Lauer-Fields J. L., Fields G. B., Visse R. and Nagase H., *EMBO J.*, 2004, 23, 3020-3030

32. Aimes R. T., Quigley J. P., *J. Biol. Chem*., 1995, 270, 5872-5876

33. Boire A., Covic L., Agarwal A., Jacques S., Sherifi S., and Kuliopulos A., *Cell*, 2005, 120, 303–313

34. Murphy G., Nguyen Q., Cockett M. I., Atkinson S. J., Allan J. A., *J. Biol. Chem.*, 1994, 269, 6632–6636

35. Shipley J. M., Doyle G. A., Fliszar C. J., Ye Q. Z., Johnson L. L., *J. Biol. Chem.*, 1996, 271, 4335–4341

36. Strongin A. Y., Collier I., Bannikov G., Marmer B. L., Grant G. A. and Goldberg G. I. *J. Biol. Chem.,* 1995, 270, 5331-5338

37. Crabbe J. C., Wahlsten D., Dudek B. C., *Science,* 1996, 284, 1667-1670

38. Murphy G., Segain J. P., O'Shea M., Cockett M., Ioannou C., Lefebvre O., *J. Biol. Chem*., 1993, 268, 15435-15441

39. Marchenko N. D., Marchenko G. N., Weinreb R. N., Lindsey J. D., Kyshtoobayeva A., Crawford H. C., *Int. J. Biochem. Cell Biol*., 2004, 36, 942-956

40. Ohuchi E., Imai K., Fujii Y., Sato H., Seiki M., Okada Y., *J. Biol. Chem.,* 1997, 272, 2446-2451

41. Shapiro S. D., Kobayashi D. K. and Ley T. J., *J. Biol. Chem.,* 1993, 268, 23824-23829

42. Gronski T. J. Jr, Martin R. L., Kobayashi D. K., Walsh B. C., Holman M. C., Huber M., Van Wart H. E. and Shapiro S. D., *J. Biol. Chem.,* 1997, 272, 12189-12194

43. Chandler S., Cossins J., Lury J. and Wells G., *Biochem. Biophys. Res. Commun.*, 1996, 228, 421-429

44. Shipley J. M., Wesselschmidt R. L., Kobayashi D. K., Ley T. J., Shapiro S. D., *Proc. Natl. Acad. Sci. U S A,* 1996, 93, 3942-1946

45. Stracke J. O,, Hutton M,, Stewart M,, Pendas A, M,, Smith B. and Lopez-Otin C,, *J. Biol. Chem*., 2000, 275, 14809-14816

46. Sedlacek R., Mauch S., Kolb B., Schatzlein C., Eibel H., Peter H. H., *Immunobiology,* 1998, 198, 408-423

47. Ryu O. H., Fincham A. G., Hu C. C., Zhang C., Qian Q. and Bartlett J. D., *J. Dent. Res.*, 1999, 78, 743-750

48. Yang M. Z., Murray M. T., Kurkinen M., *J. Bio.l Chem.* 1997, 272, 13527-13533

49. Marchenko G. N., Marchenko N. D., Strongin A. Y., *Biochem. J*. 2003, 372, 503-515

50. Ahokas K., Lohi J., Illman S. A., Llano E., Elomaa O., Impola U. *Lab Invest*, 2003, 83, 1887-1899

51. Velasco G., Pendas A. M., Fueyo A., Knäuper V., Murphy G., Lo´pez-Otin C., *J. Biol. Chem*., 1999, 274, 4570-4576

52. Yang M. Z., Kurkinen M., *J. Bio.l Chem.*, 1998, 273, 17893-17900

53. Saarialho-Kere U., Kerkela E., Jahkola T., Suomela S., Keski-Oja J., Lohi J., *J. Invest. Dermatol*. 2002, 119, 14-21

54. Rasmussen H. S. and McCann P. P., *Pharmacol. Ther.*, 1997, 75, 69-75

55. Overall C. M. and López-Otín C., *Nat. Rev. Cancer*, 2002, 2, 657-672

56. Goldberg G. I., Marmer B. L., Grant, G. A. Eisen A. Z., Wilhelm S. and He C. S., *Proc. Natl. Acad. Sci.* 1989, 86. 8207-8211

57. Gomez D. E., Alonso D. F., Yoshiji H., Thorgeirsson U. P., *Eur. J. Cell. Biol*. 1997, 74, 111-122

58. Dhanaraj V., Ye Q. Z., Johnson L. L., Hupe D. J., Ortwine D. F., Dunbar J. B. Jr., Rubin. J. R., Pavlovsky A., Humblet C. and Blundell T. L., *Structure*, 1996, 4, 375-386

59. Park H. I., Jin Y., Hurst D. R., Monroe C. A., Lee S., Schwartz M. A. and Sang Q. X., *J. Biol. Chem*., 2003, 278, 51646-51653

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

61. Andreini C., Banci L., Bertini I., Luchinat C. and Rosato A., *J. Proteome Res.*, 2004, 3, 21-31

62. Murphy G., Allan J. A., Willenbrock F., Cockett M. I., O'connell J. P. and Docherty A. J., *J. Biol. Chem.*, 1992, 267, 9612-9618

63. Knäupper V., Cowell S., Smith B., Lopez-Otin C., O'Shea M., Morris H., Zardi L. Murphy G., *J. Biol. Chem*., 1997, 272, 7608-7616

64. Chung et al, *EMBO J*., 2004, 23, 3020-3030

65. Bode W., Reinemerl P., Huber R., Kleine T., Schnierer S. and Tschesche H., *EMBO J*., 1994, 13, 1263-1269

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

67. Cha J. and Auld D. S., *Biochemistry,* 1997, *36,* 16019-16024

68. Tam E. M., Wu Y. I., Butler G. S., Stack M. S., and Overall C. M. 2002, *J. Biol. Chem.* 277, 39005–39014.

69. Osenkowski P., Meroueh S. O., Pavel D., Mobashery S., and Fridman R., *J. Biol. Chem.,* 2005, 280, 26160–26168

70. Knauper V., Docherty A. J. P., Smith B., Tschesche H., Murphy G., *FEBS Letters,* 1997, 405 60-64

71. Knauper V., Osthues A., Declerck Y. A., Langley K. E., Blaser J., and Tschesche H., *Biochem. J.,* 1993, 291, 847-854

72. Cha J., Pedersen M. V. and Auld D. S. *Biochemistry* 1996**,** *35,* 15831-15838

73. Barnett J., Straub K., Nguyen B., Chow J., Suttman R., Thompson K., Tsing S., Benton P., Schatzman R., Chen M., and Chan H. *Protein Expr. Purif.,* 1994, 5, 27-36