

Matrilysin–Inhibitor Complexes: Common Themes among Metalloproteases[‡]Michelle F. Browner,^{*,§} Ward W. Smith,^{||,#} and Arlindo L. Castelhanol[‡]

Molecular Structure Department and Institute of Organic Chemistry, Syntex Discovery Research, 3401 Hillview Avenue, Palo Alto, California 94303, and Agouron Pharmaceuticals, Inc., 3565 General Atomics Court, San Diego, California 92121

Received November 4, 1994; Revised Manuscript Received February 15, 1995[®]

ABSTRACT: Matrix metalloproteases are a family of enzymes that play critical roles in the physiological and pathological degradation of the extracellular matrix. These enzymes may be important therapeutic targets for the treatment of various diseases where tissue degradation is part of the pathology, such as cancer and arthritis. Matrilysin is the smallest member of this family of enzymes, all of which require zinc for catalytic activity. The first X-ray crystal structures of human matrilysin are presented. Inhibitors of metalloproteases are often characterized by the chemical group that interacts with the active site zinc of the protein. The structures of matrilysin complexed with hydroxamate (maximum resolution 1.9 Å), carboxylate (maximum resolution 2.4 Å), and sulfodiimine (maximum resolution 2.3 Å) inhibitors are presented here and provide detailed information about how each functional group interacts with the catalytic zinc. Only the zinc-coordination group is variable in this series of inhibitors. Examination of these inhibitor–matrilysin complexes emphasizes the dominant role the zinc-coordinating group plays in determining the relative potencies of the inhibitors. The structures of these matrilysin–inhibitor complexes also provide a basis for comparing the catalytic mechanism of MMPs and other metalloproteins.

Matrix metalloproteinases (MMPs)¹ play a prominent role in the degradation of extracellular matrix during tissue morphogenesis, differentiation, and wound healing (Birkedal-Hansen et al., 1993). MMPs have also been implicated in various pathologies, including tumor invasion and metastasis (Stetler-Stevenson et al., 1993). Components of the extracellular matrix subject to degradation by MMPs include fibrillar collagens, proteoglycans, fibronectin, and gelatin. Physiological and cellular regulation of MMP activities is not well understood, but a complex array of cellular factors, such as oncogenes, growth factors, hormones, and specific protein inhibitors, is involved in controlling expression and latency (Matrisian, 1992).

The broad substrate specificity required to degrade all components of the extracellular matrix is achieved within the MMP family through evolutionary redundancy and divergence. To date, at least 11 different human enzymes have been identified and are classified as either collagenases, gelatinases, or stromelysins, on the basis of their general substrate specificities (Birkedal-Hansen et al., 1993). Matrilysin is most similar to stromelysin when amino acid sequences are compared.

All MMPs so far described share an obvious primary domain structure that has been characterized biochemically and by amino acid sequence comparisons. The N-terminal domain, known as the pro domain, is present in all MMPs and is responsible for the latency of the zymogen. A cysteine residue present in the pro domain has been proposed to coordinate the active site zinc (Springman et al., 1990; VanWart & Birkedal-Hansen, 1990), and this arrangement has now been confirmed by structural studies and biochemical studies (K. Appelt, W. Smith, and R. Almasy, in preparation). The N-terminal domain is proteolytically cleaved from the catalytic domain to produce the active enzyme. The MMP catalytic domain is characterized by a conserved sequence motif (HEXGHXXGXXHS) responsible for coordinating bound zinc and containing the glutamate residue that is critical for catalysis.

All MMP enzymes, except matrilysin, have a C-terminal domain that has amino acid sequence similarity to hemopexin, a plasma heme binding protein. This domain has been shown to have a role in defining the substrate specificity of some MMPs. This is most dramatic in the case of the collagenase subfamily; in the absence of the C-terminal domain, triple-helical collagen is no longer an *in vitro* substrate for these enzymes (Clark & Cawston, 1989; Murphy et al., 1992; Sanchez-Lopez et al., 1993; Windsor et al., 1991). On the other hand, the substrate specificity of matrilysin is only dependent on the substrate binding subsites (S₁, S₁', etc.) in its single domain. Matrilysin shows good enzymatic activity against a wide range of peptide substrates. Matrilysin can efficiently hydrolyze peptides with many different P₁ residues including alanine, methionine, and proline, which are preferred by the collagenases, and tyrosine, which is favored by the gelatinase enzymes (Netzel-Arnett et al., 1991, 1993).

Recently, the X-ray crystal structures of two different mature, truncated (no C-terminal domain) forms of collagenase (Bode et al., 1994; Borkakoti et al., 1994; Lovejoy et

[‡] Coordinates have been deposited in the Brookhaven Protein Data Bank under the file names 1MMP, 1MMQ, and 1MMR.

^{*} Corresponding author [telephone, (415) 354-2238; fax, (415) 354-7363; email, browner@arvax.syntex.com].

[§] Molecular Structure Department, Syntex Discovery Research.

^{||} Agouron Pharmaceuticals, Inc.

[#] Current address: Department of Macromolecular Sciences, Smith-Kline Beecham Pharmaceuticals, 709 Swedeland Road, UE0447, King of Prussia, PA 19406.

¹ Institute of Organic Chemistry, Syntex Discovery Research.

[®] Abstract published in *Advance ACS Abstracts*, April 1, 1995.

¹ Abbreviations: MMP, matrix metalloproteases; Tris-HCl, Tris-(hydroxymethyl)aminomethane hydrochloride; BES, N,N-bis(2-hydroxyethyl)-2-aminoethanesulfonic acid; DNP, 2,4-dinitrophenyl; HEPES, N-(2-hydroxyethyl)piperazine-N'-2-ethanesulfonic acid; DMSO, dimethyl sulfoxide.

al., 1994a; Spurlino et al., 1994; Stams et al., 1994) bound with inhibitors have been reported. The structure of mature, truncated stromelysin as determined by NMR spectroscopy has also been reported (Gooley et al., 1994). This paper reports the first X-ray crystal structures of mature matrilysin in complex with several different inhibitors.

Much experimental work has been done to characterize the catalytic mechanisms of other zinc-dependent proteases, such as thermolysin and carboxypeptidase A. X-ray crystallographic studies of various inhibitor–protein complexes with different zinc ligands have provided plausible models for how these enzymes interact with substrates and for the configuration of the transition state (Christianson & Lipscomb, 1989; Matthews, 1988). This study provides the framework from which a similarly detailed understanding of the catalytic mechanism of matrix metalloproteinases can be obtained.

EXPERIMENTAL PROCEDURES

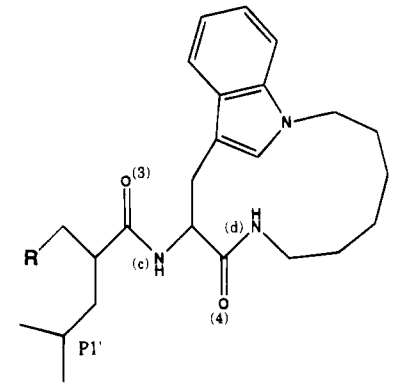
Protein Purification. Recombinant human matrilysin protein was obtained from a mammalian cell culture expression system (Barnett et al., 1994). The zymogen is secreted from Chinese hamster ovary cells and then purified by sequential ion-exchange and metal chelation chromatography. To obtain high yields of protein from this procedure, the detergent Brij 35 (0.005%) was present throughout the purification. Initial crystallization trials using either zymogen or activated protein obtained directly from this procedure were not promising.

Recombinant zymogen protein was dialyzed into buffer A (20 mM BES, pH 7.0, 5 mM CaCl_2 , 150 mM NaCl) without Brij 35; detergent, however, was still present as shown by NMR spectroscopy. Complete removal of the detergent was effected using a modified metal chelation chromatography procedure: chelating Superose HR (Pharmacia) charged with CuCl_2 . The matrilysin zymogen was bound in buffer A to the metal chelate and released by increasing the concentration of imidazole (0–100 mM) in buffer A. The zymogen was activated with *p*-aminophenylmercuric acetate and assayed as described previously (Barnett et al., 1994). Mature matrilysin was dialyzed into 20 mM Tris-HCl, pH 8.0, 5 mM CaCl_2 , and 5 mM NaCl, concentrated, and stored at -70°C .

Enzyme activity was determined by a continuous fluorescence assay that utilizes the fluorogenic heptapeptide substrate 2,4-DNP-Pro-Leu-Gly-Leu-Trp-Ala-D-Arg-NH₂ (Barnett et al., 1994; Knight et al., 1992). The hydroxamate and carboxylate inhibitors were synthesized by A. Castelhana and co-workers (in preparation), and the sulfodiimine compound was prepared in collaboration with M. Schwartz (in preparation) at Florida State University. Effective inhibitor concentrations were determined using the same assay with various inhibitor concentrations.

Crystallization. Crystals of each protein–inhibitor complex were obtained by screening crystallization conditions (Jancarik & Kim, 1991) with the inhibitor (hydroxamate, carboxylate, or sulfodiimine; Table 1) present at twice the molar protein concentration. In the crystallization drops, the final concentrations of protein and inhibitor were 10 mg/mL and 1.2 mM (in 2.5% DMSO), respectively. Crystals were obtained by vapor diffusion using the hanging drop method (McPherson, 1982) at 17 – 25°C . Matrilysin crystals

Table 1: Matrilysin Inhibitors

			
R	hydroxamate	carboxylate	sulfodiimine
Ki	0.03 μM	0.85 μM	4.0 μM

in the presence of the hydroxamate inhibitor appeared when the well solution contained 0.2 M CaCl_2 , 0.1 M HEPES, pH 7.5, and 30% 2-propanol. Crystals in the presence of the carboxylate inhibitor grew when the well solution contained 0.2 M sodium citrate, 0.1 M cacodylate, pH 6.5, and 30% 2-propanol. Crystallization of matrilysin with the sulfodiimine inhibitor required 1.5 M NaCl and 10% ethanol in the well solution. Crystals typically appeared within 10 days.

X-ray Diffraction Data and Structure Determination. All data were collected on a MAR (300-mm diameter) image plate detector using Cu $K\alpha$ radiation reflected from a graphite monochromator. The data were reduced using MOSFLM (version 5.1.2) and scaled using ROTAVATA and AGROVATA (CCP4, 1994). A summary of the X-ray diffraction data is presented in Table 2.

The molecular replacement method (Rossman, 1990) was used to obtain the initial, approximate phase information for the hydroxamate and carboxylate protein complexes. The molecular replacement model for the hydroxamate complex was derived from a partial X-ray crystal structure model of the prostromelysin catalytic domain (K. Appelt, W. Smith, and R. Almasy, private communication). The model included all main-chain atoms for the amino acid residues of the catalytic domain (100–239 and 250–267) and the side-chain atoms of amino acid residues that are identical between matrilysin and stromelysin (94 out of 161 amino acid residues). A single, unambiguous solution to the rotation search was obtained when the initial top 1000 peaks were refined by Patterson correlation, using several different resolution ranges (Brunger, 1987b, 1990). The translation search was solved (top peak 17σ) only in space group $P3_121$; after rigid body refinement the R_{cryst} (see Table 2, footnote b) was 45.8% for data from 7.0 to 3.0 Å. Clear features present in the $2F_o - F_c$ and $F_o - F_c$ electron density maps indicated the correctness of the solution, including several bound metals and well-shaped electron density for much of the hydroxamate inhibitor. Positional refinement by molecular dynamics (Brunger, 1987a, 1987b) provided a starting model for matrilysin ($R_{\text{cryst}} = 37.5\%$, 8.0–2.3 Å) that was modified using the program CHAIN (Sack, 1988).

Table 2: Matrilysin–Inhibitor Complexes: Structure Determination Summary

	hydroxamate	carboxylate	sulfodiimine
symmetry	$P3_121$	$P3_121$	$P3_121$
unit cell	$a = b = 61.9 \text{ \AA}$, $c = 88.0 \text{ \AA}$, $\gamma = 120^\circ$	$a = b = 62.0 \text{ \AA}$, $c = 175.7 \text{ \AA}$, $\gamma = 120^\circ$	$a = b = 62.1 \text{ \AA}$, $c = 87.4 \text{ \AA}$, $\gamma = 120^\circ$
asymmetric unit	monomer	dimer	monomer
maximum resolution (\AA)	1.9	2.4	2.3
total reflections	83 992	80 416	23 346
unique reflections	15 546	15 769	7 690
average I/σ	5.0	7.4	7.2
R_{sym}^a	0.12	0.09	0.08
structure solution	MR, stromelysin model	MR, matrilysin model	X-PLOR
refinement	X-PLOR	X-PLOR	X-PLOR
resolution (\AA)	6.0–1.9	6.0–2.4	6.0–2.4
R_{cryst}^b (%)	17.9	18.0	18.7
nonhydrogen atoms	1322 + 70 H_2O	2620 + 108 H_2O	1311 + 57 H_2O
RMS bonds (\AA)	0.013	0.013	0.014
RMS angles (deg)	2.73	2.76	2.88

^a $R_{\text{sym}} = \sum_h \sum_i |\langle F_h \rangle - F_{hi}| / \sum_h F_h$, where $\langle F_h \rangle$ is the mean structure factor amplitude of the i observations of reflections that are related to the Bragg index h . ^b $R_{\text{cryst}} = \sum ||F_o| - |F_c|| / \sum |F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively.

After rebuilding the protein chain and adding four metals and the hydroxamate inhibitor, a cycle of refinement by simulated annealing (Brunger, 1987b; Kirkpatrick et al., 1983) reduced the R_{cryst} significantly (from 28.2% to 25.3% for the 6.0–1.9 \AA data). Further iterative cycles of computational and graphical refinement of the model provided a final structure with good geometry and an R_{cryst} of 17.9% (Table 2).

The structure of the matrilysin–carboxylate inhibitor complex was determined using the structure of the hydroxamate complex as a starting model without the bound metals and inhibitor. A single peak was identified in the rotation search using X-PLOR. A full cell translation search using the program BRUTE was carried out with the single rotation solution (Fujinaga & Read, 1987). The top peak in space group $P3_121$ was 8.9σ with the next highest peak at 7.2σ . In space group $P3_121$ the best solution was 5.6σ . Inspection of a native Patterson map indicated that there was a noncrystallographic 2-fold along the z -axis at (0, 0, 0.25). The initial dimer model was obtained by applying the noncrystallographic 2-fold to the solution determined in space group $P3_121$. The two monomers contained in the asymmetric unit of this crystal form are related by a rotation of 178.5° about a noncrystallographic 2-fold axis. The electron density map calculated with the initial model after rigid body refinement ($R_{\text{cryst}} = 30.8\%$ for data between 8.0 and 3.0 \AA resolution) clearly showed the bound metals and inhibitor. The refinement process was essentially as described above, except that atoms related by noncrystallographic symmetry were restrained to their average positions (0.2 \AA RMS deviation), as implemented in XPLOR (Brunger, 1987b) (Table 2).

Although the crystallization conditions for the matrilysin–sulfodiimine complex were not similar to the crystallization parameters used for the hydroxamate complex, the crystals are isomorphous. The first electron density map for the sulfodiimine complex was calculated using the final hydroxamate structure without inhibitor as the starting model ($R_{\text{cryst}} = 44.0\%$ for all data between 8.0 and 2.8 \AA). Rigid body refinement gave a model with an $R_{\text{cryst}} = 32.4\%$. Clear electron density allowed accurate positioning of the sulfodiimine inhibitor, and the complex was refined as discussed previously (Table 2).

X-PLOR parameters, including bond and angle connectivities as appropriate for zinc–histidine interactions, were

implemented as outlined previously for protein-bound copper (McGrath et al., 1993). For calcium, only nonbonded parameters were included; these parameters allow short distances between the calcium atom and its ligands. X-PLOR parameters for each inhibitor were implemented using atom definitions and parameters developed for X-PLOR by Molecular Simulations, Inc.

The numbering of amino acid residues for matrilysin in the structures presented begins at 100, for the N-terminal residue (tyrosine) of the mature enzyme. This numbering corresponds to the numbering used in two collagenase structures (Spurlino et al., 1994; Stams et al., 1994).

RESULTS

General Structure Description. The X-ray crystal structure of matrilysin shows that the protein fold is an open-face, α – β sandwich (Figure 1). The overall fold is the same as that reported for other members of the MMP family (Bode et al., 1994; Borkakoti et al., 1994; Gooley et al., 1994; Lovejoy et al., 1994a; Spurlino et al., 1994; Stams et al., 1994). The five-stranded β -sheet is formed by four parallel strands (1, 2, 3, and 5) and one antiparallel β -strand (4). The long N-terminal α -helix (A) runs along the line of β -strand 2, which is twisted relative to the β -sheet. The central helix (B) lies under the β -sheet, at approximately a 20° rotation relative to helix A. The third helix (C) lies on the outer surface, orthogonal to helix A. The topological equivalence of the MMP catalytic domains is not surprising since pairwise combinations comparing matrilysin, stromelysin, and the collagenase enzymes reveal 49–56% amino acid identity. Like other MMPs, matrilysin is also structurally similar to less closely related metalloproteases such as astacin, adamalysin, and thermolysin (Bode et al., 1993; Borkakoti et al., 1994; Hodgkin et al., 1994; Monzingo & Matthews, 1984; Spurlino et al., 1994).

The N-terminal peptide of mature matrilysin has a well-defined binding site on the outside of the catalytic domain. The amino-terminal group forms a salt bridge with the carboxylate of the conserved residue Asp254; the distances from the terminal nitrogen to the carboxylate oxygens are 2.6 and 3.1 \AA . A very similar position for the N-terminal peptide of one of the collagenase subfamily enzymes was recently reported (Reinemer et al., 1994). A short, twisted, two-stranded β -sheet is formed between N-terminal residues

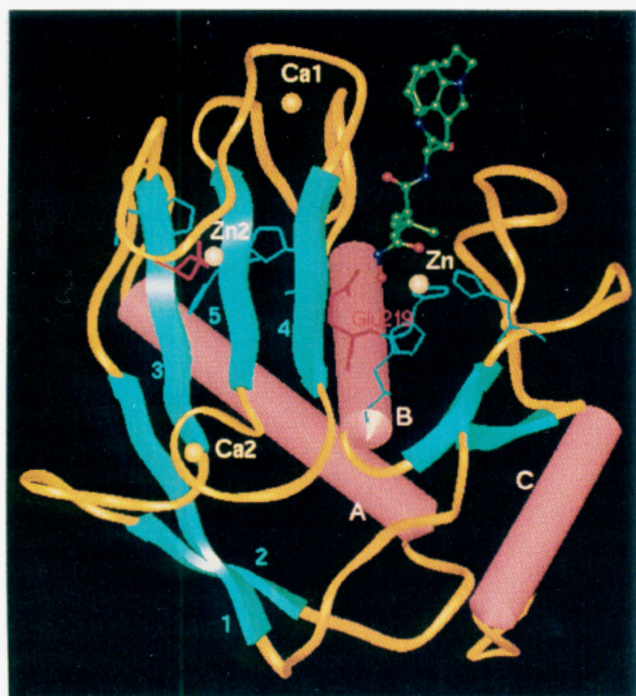


FIGURE 1: Secondary structure diagram of matrilysin with the hydroxamate inhibitor bound. α -Helices (labeled A–C) are represented as pink cylinders, and β -strands (numbered 1–5) are shown as blue arrows. The short β -sheet formed by the binding site for the N-terminus of mature matrilysin is also shown. The four bound metals are represented by spheres.

101–103 and residues 225–227, which are C-terminal to helix B (Figure 1). The hydroxyl group of Tyr100 (the amino-terminal residue) is within hydrogen-bonding distance of Lys257 (which is just before helix C), further stabilizing the position of the N-terminal peptide in matrilysin. Presumably, activation of the zymogen involves movement of the N-terminal peptide to this binding site.

Metal Binding Sites. The active site of matrilysin lies in the relatively shallow cleft between the β -sheet and the central helix (B) and is occupied by the inhibitor in each

complex (Figures 1 and 2). The catalytic zinc is coordinated by the N ϵ of His218, -222, and -228; the average N ϵ to Zn distance for all three matrilysin–inhibitor complexes is 2.2 Å. Other zinc ligands are supplied by the atoms from the bound inhibitors (discussed below).

Other than the catalytic zinc, three other metal ions, one zinc ion and two calcium ions, are bound in these X-ray crystal structures of matrilysin (Figure 1). Together these three metal ions appear to position the antiparallel β -sheet that forms one side of the active site cleft. The structural zinc (Zn2) is tetracoordinate, bound by the atoms His168-N ϵ and Glu170-O ϵ from the loop/turn that connects strands 3 and 4. The other two zinc ligands are His183-N ϵ from strand 5 and His196-N δ from strand 4. Further stabilization of the five-stranded β -sheet is provided by the two calcium ions. Identification of the calcium ions was based on the presence of CaCl₂ in all protein solutions and the observed coordination geometry. Both calcium ions are hexacoordinate, and the range of ligand–calcium distance is 2.4–2.1 Å. Calcium 1 (Ca1) is coordinated by six protein atoms. Four atoms in the long loop between strands 3 and 4, closer to strand 4, bind Ca1: Glu175 and the carbonyl oxygens of residues 176, 178, and 180. The other two ligands to Ca1 are on the loop connecting strand 5 and helix B (Glu198 and Glu201). The second calcium is coordinated by only four protein atoms: the carbonyl oxygen of residue 158 located in the loop between strands 2 and 3, carbonyl oxygens of residues 190 and 192 between strands 4 and 5, and the O δ of Asp178 in strand 5. The other two ligands are water molecules that are within hydrogen-bonding distance of protein atoms. The positions of four strands of the five-stranded β -sheet are stabilized by the bound metal ions, both at the top and bottom of the sheet. Biochemical experiments have shown that, in addition to the catalytic requirement for zinc, MMPs are stabilized by the presence of both zinc and calcium (Housely et al., 1993; Lowry et al., 1992; Okada et al., 1986).

Inhibitor Binding Site. The inhibitors used in this investigation were the result of a medicinal chemistry program

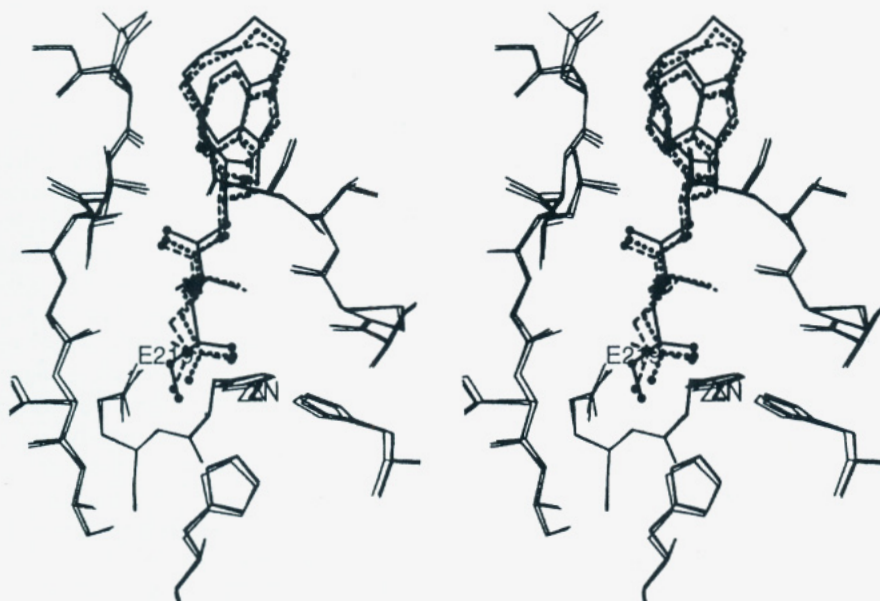


FIGURE 2: Stereo diagram of the inhibitor binding site. The protein backbone atoms from each inhibitor complex are superimposed, and the positional overlap of the bound inhibitors is shown; the hydroxamate compound is shown by solid lines, the carboxylate compound by dotted lines, and the sulfodiimine compound by dashed lines. The heteroatoms of the inhibitors are shown as solid circles.

Table 3: Polar Interactions between Matrilysin and Bound Inhibitors

protein atom	hydroxamate		carboxylate		sulfodiimine	
	atom ^a	distance (Å)	atom	distance (Å)	atom	distance (Å)
Zn	O ₍₁₎	2.2	O ₍₁₎	2.1	N _(a)	2.0
	O ₍₂₎	2.2	O ₍₂₎	2.4	N _(b)	2.9
Glu203-OE1	O ₍₂₎	2.6	O ₍₂₎	3.3	N _(b)	3.4
	N _(a)	3.0				
Glu203-OE2	O ₍₂₎	3.0	O ₍₂₎	2.7	N _(b)	2.9
Ala166-O	N _(a)	2.9				
Leu165-N	O ₍₃₎	2.9	O ₍₃₎	2.84	O ₍₃₎	2.8
Pro222-O	N _(c)	3.0	N _(c)	3.1	N _(c)	3.0
Tyr224-N	O ₍₄₎	2.8	O ₍₄₎	2.8	O ₍₄₎	2.8
Asn163-O	N _(d)	2.9	N _(d)	2.9	N _(d)	2.8
H ₂ O ₍₁₎	O ₍₁₎	2.7	O ₍₁₎	2.9	N _(a)	2.6
H ₂ O ₍₂₎	O ₍₂₎	2.7	O ₍₂₎	2.7	N _(b)	2.8

^a Numbering of inhibitor atoms is presented in Figure 1.

involving substrate analogs, initiated before structure information for any MMP was available. Each inhibitor in this series has a unique zinc-coordinating ligand (hydroxamate, carboxylate, or sulfodiimine; Table 1). The remainder of each inhibitor is identical and includes a leucine group at the position equivalent to the P₁' residue of a peptide substrate and a 13-member macrocycle, which includes an indole-lactam ring in the position analogous to the P₂' residue of a peptide substrate. The hydroxamate and carboxylate inhibitors are bound in virtually identical conformations, with an RMS deviation of 0.2 Å for all atoms excluding the zinc ligand and the adjacent methylene (Table 1 and Figure 2). The sulfodiimine inhibitor has adopted a slightly different conformation with an RMS deviation of 0.5 Å compared to the hydroxamate. The larger RMS deviation for the sulfodiimine compound is due to a positional difference in the six-carbon methylene chain of the macrocycle.

All of the polar atoms in the three inhibitors interact with main-chain atoms of matrilysin. Table 3 shows the distances between atoms of each bound inhibitor and the protein. All common interactions between the inhibitor and protein atoms (middle portion of Table 3) are identical, within experimental error. The calculated RMS deviations relating the protein main-chain atoms of the hydroxamate complex with those of the carboxylate and sulfodiimine complexes are both 0.2 Å. The positions of these atoms can therefore be considered to be identical since 0.2 Å is the estimated error on the atomic coordinates (Luzzati, 1952). There are, however, significant positional differences between the bound inhibitors (Figure 2). With protein backbone atoms used in superimposing the structures, there are significant variations in the positions of the polar atoms of the inhibitors (a mean variation of 0.5 Å). For example, the position of the second inhibitor carbonyl oxygen (O₍₄₎; Table 1) is displaced by 0.7 Å in the carboxylate and sulfodiimine complexes, relative to its position in the hydroxamate structure, and both amino groups (N_(c) and N_(d)) show a positional variation of 0.5 Å (Figure 2).

There is also a large positional variation of the methylene that is bonded to the zinc liganding group of each inhibitor. The positional difference for this methylene is 1.5 Å for the sulfodiimine complex compared to the hydroxamate structure. The methylene of the sulfodiimine compound projects furthest into the active site cleft, toward β -strand 4, whereas the methylene of the hydroxamate protrudes out of the active

site cleft. The same methylene in the carboxylate structure is positioned in between, displaced 0.7 Å when compared to the hydroxamate or sulfodiimine complex. It is at the position of this methylene that each inhibitor adopts a unique configuration in its interaction with the active site zinc.

The S₁' pocket in matrilysin is defined by three aromatic side chains, Tyr214, His218, and Tyr240. Tyr214 is unique to matrilysin and lies at the back of the S₁' pocket. The P₁' residue of the bound inhibitors is a leucine, which is a good residue for this position in peptide substrates, as determined by the rate of peptide hydrolysis (Netzel-Arnett et al., 1993). Leucine fills the S₁' pocket adequately, although there is clearly room for "bulkier" side chains such as isoleucine, which is the preferred P₁' residue for substrates. This is highlighted by the sulfodiimine structure, in which one of the δ -methyl groups of the P₁' residue protrudes deeper (by 0.8 Å) into the S₁' pocket. There is no difference in the position of the protein atoms that line the S₁' pocket in response to the shift of the P₁' residue in the sulfodiimine complex. In agreement with the P₁' substrate specificity established by kinetic analysis (Netzel-Arnett et al., 1993), the S₁' pocket of matrilysin seems to accept positionally variant hydrophobic groups of different shapes and volumes.

The indole-lactam group of the inhibitor is in a position presumably analogous to the P₂' residue of a peptide substrate. It does not, however, appear to bind in a well-defined S₂' pocket of the enzyme (Figure 2). The indole-lactam ring is interacting with Asn179 through van der Waals contacts; Asn179 is on the surface of the enzyme, in the loop between strands 3 and 4 where Ca1 is bound. It is likely that the binding site of the indole-lactam ring is influenced by the methylene linker and, therefore, does not truly represent a P₂' residue of a peptide substrate. The methylene chain of the macrocycle extends into the active site cleft.

Active Site Zinc-Inhibitor Interactions. The configuration and bonding interactions for the zinc liganding group of each inhibitor are unique. The hydroxamate group provides a perfect bidentate ligand to the zinc as reported previously for thermolysin (Holmes & Matthews, 1981); each oxygen of the hydroxamate group is 2.2 Å from the active site zinc. In the matrilysin-hydroxamate complex, the nitrogen is in position to interact with the carbonyl oxygen of Ala182 (Table 3 and Figure 3), suggesting that the hydroxamate nitrogen is protonated. This observation was confirmed by NMR spectroscopy using a ¹⁵N-enriched hydroxamate inhibitor (J. Pease and S. Bender, private communication). In the thermolysin-hydroxamate complex, however, the hydroxamate nitrogen was not positioned to form a hydrogen bond with a protein atom. This observation led to the conclusion that the anionic, N-acid, form of the hydroxamate is bound in thermolysin (Holmes & Matthews, 1981). Data from the X-ray crystal structure, and supported by NMR experiments, clearly show that the O-acid form of the hydroxamate is bound to matrilysin.

The interaction of the carboxylate group with the active site zinc creates a quasi-pentacoordinate zinc-ligand system. The carboxylate oxygens are at unequal distances to the zinc, but both are still within the inner zinc coordination sphere (Table 3). The syn stereochemistry observed for the carboxylate interaction is the preferred configuration (Christianson, 1991). The oxygen-zinc distances are remarkably similar to those reported for a thermolysin-carboxylate

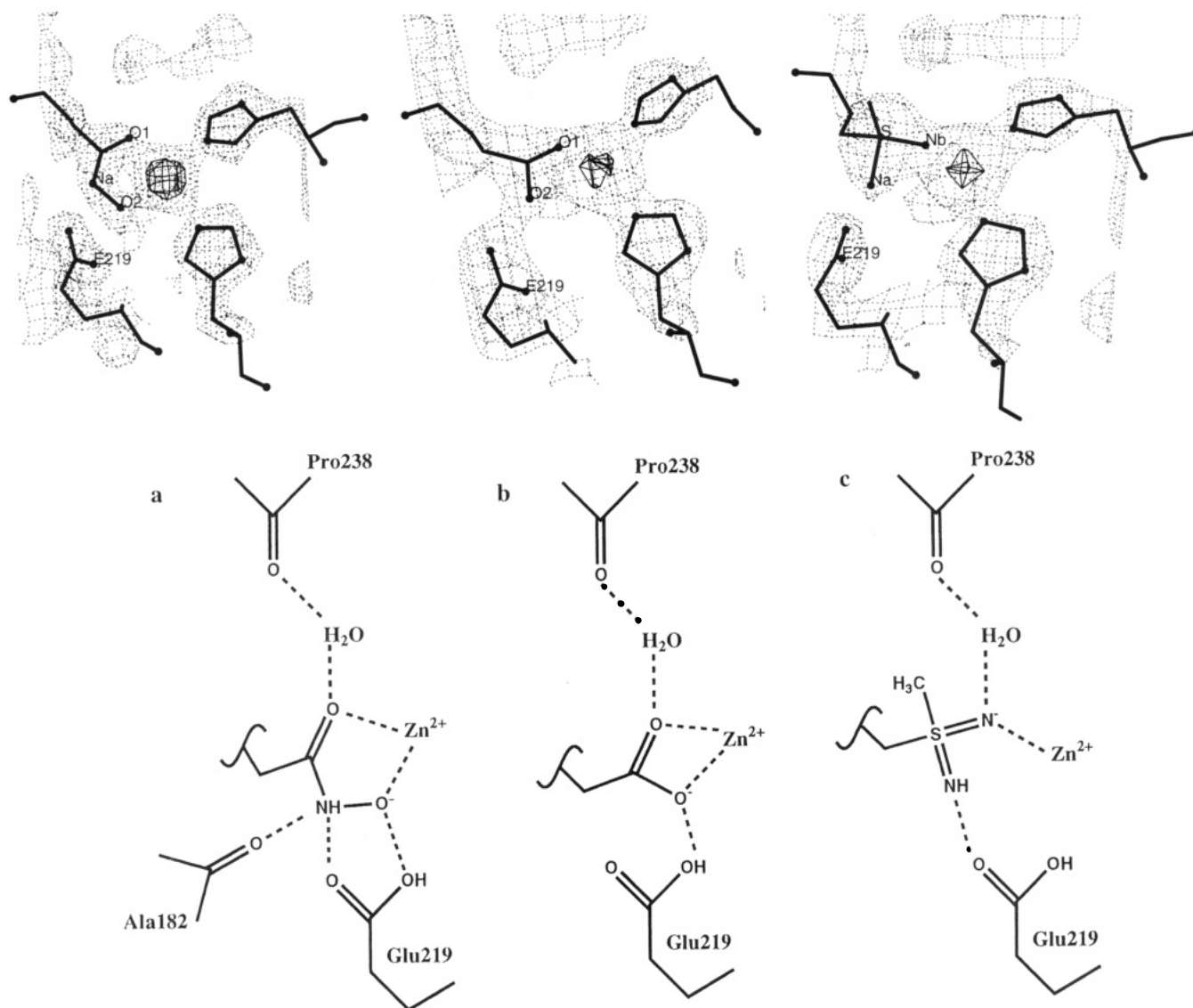


FIGURE 3: Inhibitor interactions with the catalytic zinc. (A, top) $2F_o - F_c$ electron density maps in the region of the active site zinc. Two of the histidines that are zinc ligands from the protein are shown. The position of the zinc is indicated by the solid line electron density (10σ). The atoms of the inhibitor liganding groups are labeled: (1) hydroxamate inhibitor, 2σ electron density map; (2) carboxylate inhibitor, 1.5σ electron density map; and (3) sulfodiimine inhibitor, 1.5σ electron density map. (B, bottom) Schematic diagram of the interactions of each inhibitor with the active site zinc and the conserved water molecule. a, b, and c as in part A. Only the zinc ligand of each inhibitor is represented.

inhibitor complex (2.0 and 2.4 Å) (Monzingo & Matthews, 1984). In both the hydroxamate- and carboxylate-bound protein structures, there are two water molecules near the active site zinc that are in position to interact with inhibitor–ligand groups (Table 3).

Unlike either the hydroxamate or carboxylate inhibitor interactions, the sulfodiimine–zinc interaction is clearly a unidentate interaction in the matrilysin–inhibitor complex. One nitrogen is 2.0 Å from the zinc and clearly makes a strong interaction with the ion. The second nitrogen, at a distance of 2.9 Å, is too distant to be an inner-sphere zinc ligand. A similar, unidentate mode of zinc interaction was also observed for a sulfodiimine bound to carboxypeptidase A (Cappalonga et al., 1992). There are, however, significant differences between the sulfodiimine configuration in the matrilysin and carboxypeptidase A structures (Figure 4). The sulfodiimine bound to matrilysin is in a gauche conformation, with the imino group ($N_{(a)}$; Table 3) which is not ligating zinc positioned to interact with the catalytic residue Glu219. The imino group ($N_{(b)}$) that is a ligand for the matrilysin

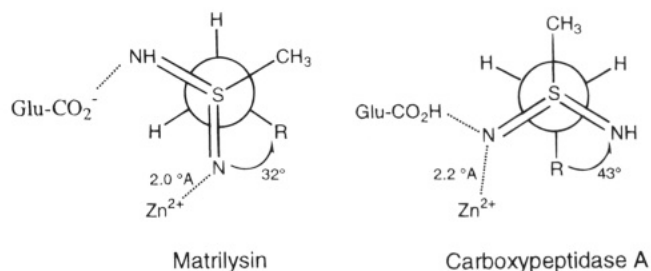


FIGURE 4: Newman projections illustrating the conformation of the sulfodiimine group when complexed with matrilysin and carboxypeptidase A [CPS1 from the Protein Data Bank (Abola et al., 1987; Bernstein et al., 1977)]. The projection looks down the sulfur–methylene bond, and the measured dihedral angle is along the nitrogen–sulfur–methylene–R ($C\alpha$ of the P_1' residue) bond path. The nitrogen atoms that are in position to interact with the active site zinc and the glutamic acid are indicated.

active site zinc is within hydrogen-bonding distance of a conserved water molecule (Table 3, $H_2O_{(1)}$). In contrast, in the carboxypeptidase A complex [CPS1 from the Protein

Data Bank, January 1994 (Abola et al., 1987; Bernstein et al., 1977)], the sulfodiimine inhibitor adopts an anti conformation (Figure 4), with the imino-zinc ligand in position to interact with the active site glutamic acid (Cappalonga et al., 1992). The nitrogen that is not a zinc ligand in the carboxypeptidase A structure is positioned to interact with nearby Arg127. In the matrilysin structure, the methyl bonded to the sulfodiimine group is positioned pointing up out of the active site cleft. Positioning of the methyl group in the electron density was, in the end, unambiguous. When the sulfodiimine group was modeled with the methyl group in any other position, severe conflicts with protein atoms resulted; either the methyl group was in very close contact with the active site zinc (<2.0 Å) or one of the ϵ -oxygens of Glu219 (2.9 Å).

DISCUSSION

Structural characterization of mammalian matrix metalloproteases has recently been an area of intense investigation (Bode et al., 1994; Borkakoti et al., 1994; Gooley et al., 1994; Lovejoy et al., 1994a; Spurlino et al., 1994; Stams et al., 1994). This is propelled, in part, by interest in the role of these enzymes in various diseases and the alluring notion that this family of zinc-dependent proteases will succumb to structure-based drug discovery technologies (Blundell, 1994). It is important to keep in mind that, to date, all reported structures have been of the catalytic domains of multidomain enzymes and thus provide only partial structural information about the biological target. For example, in the case of the collagenases, the C-terminal domain is required for hydrolysis of collagen (Clark & Cawston, 1989; Murphy et al., 1992; Sanchez-Lopez et al., 1993; Windsor et al., 1991). Matrilysin, the smallest member of the MMP family, is an exception. Unlike all other MMPs characterized to date, full-length matrilysin has only a pro domain and a catalytic domain, and thus, the active, holoenzyme contains only a single domain. Interestingly, the overall fold of matrilysin is virtually identical to that of the catalytic domain of the collagenase enzymes (Bode et al., 1994; Borkakoti et al., 1994; Lovejoy et al., 1994a; Spurlino et al., 1994; Stams et al., 1994), as well as to stromelysin-1 (Gooley et al., 1994; K. Appelt, W. Smith, and R. Almassy, private communication). The evolutionary absence of the C-terminal domain does not dramatically affect the overall protein structure, and therefore, the X-ray crystal structure of human matrilysin does not provide obvious clues as to where the hemopexin-like domain could interact with the catalytic domain of the multidomain enzymes.

In developing MMP inhibitors, the choice of the inhibitor-zinc ligand is critical. Comparison of three different zinc ligands in the context of the same inhibitor template provides information about the role each zinc ligand plays in determining inhibitor potency. The hydroxamate, carboxylate, and sulfodiimine inhibitors each exhibit unique modes of interaction with the active site zinc. All non-zinc, polar interactions with the protein are satisfied equally well for each inhibitor, as judged by non-hydrogen atomic distances (Table 3). To a great extent the rank order of potency (Table 1) is defined by the inhibitor-zinc interactions. It is, therefore, not surprising that the hydroxamate compound is the most potent inhibitor, as it is an ideal bidentate ligand to the zinc. The potency of the hydroxamate inhibitor is also

augmented by an additional hydrogen bond interaction with the protein; the hydroxamate nitrogen ($N_{(a)}$; Table 1) is within hydrogen-bonding distance to a protein backbone carbonyl group, which also stabilizes the electronics of the zinc interaction. The carboxylate group can also serve as a bidentate ligand to the zinc, but the configuration of the dual oxygen-zinc interaction is less ideal. The sulfodiimine, which provides only a monodentate ligand to the zinc, is 100-fold less potent than the hydroxamate inhibitor.

Comparison of inhibitors which differ only in their zinc-coordinating groups indicates the extent to which the zinc binding energy dictates the potency of each compound. The structure data show the influence of the ideal zinc ligand configuration, as observed for the hydroxamate compound, given the conservation of all other inhibitor interactions. Clearly, there is an interplay between zinc coordination and the exact orientation of the other atoms in the inhibitor. In the cases presented here the zinc interactions appear to define the orientation of the inhibitor. The non-zinc, polar interactions between the inhibitor and protein are maintained by slight adjustments in the angle of the hydrogen bond partners with protein atoms.

Zinc-dependent proteolysis requires an acid group (glutamic acid), acting as a general base, in alliance with the electrophilic zinc ion to promote bond hydrolysis. The catalytic mechanisms for zinc-dependent metalloproteases are derived, to a large extent, from structural data of various inhibitor-enzyme complexes for both thermolysin and carboxypeptidase A (Christianson & Lipscomb, 1989; Matthews, 1988). Thermolysin and carboxypeptidase are structurally unrelated enzymes but, nevertheless, share common catalytic features. MMPs are topologically similar to thermolysin (Bode et al., 1993; Borkakoti et al., 1994; Hodgkin et al., 1994) but not to carboxypeptidase A, which is an exopeptidase. Not surprisingly, matrilysin, as well as other MMPs, also shares the same minimal catalytic groups as the other zinc-dependent metalloproteases, namely, glutamic acid and zinc (Figure 5).

For both thermolysin and carboxypeptidase A, other amino acid residues have been implicated in the proposed catalytic mechanisms. For example, in thermolysin the carbonyl oxygen of Ala113 and the $O\delta$ of Asn112 have both been suggested to be important in the stabilization of the protonated, scissile nitrogen (Matthews, 1988). These two interactions are conserved in fibroblast collagenase (Spurlino et al., 1994). Only the carbonyl oxygen interaction, however, is conserved in matrilysin (Figure 5) and all other MMPs. The conservation of the same enzyme-substrate interaction between thermolysin and the MMPs is indicative of their extensive structural similarities around the substrate binding site. In carboxypeptidase A only one group ($O\epsilon$ of Glu270) is in position to hydrogen bond with the leaving amino group. The active site glutamic acid of thermolysin and the MMPs can also provide this same interaction with the protonated nitrogen of the product (Figure 5).

In both thermolysin and carboxypeptidase A there is a basic amino acid positioned to form a hydrogen bond to the incoming peptide substrate and to then stabilize the tetrahedral intermediate. His231 and Tyr157 in thermolysin and Arg127 in carboxypeptidase A (Phillips et al., 1990) have been suggested to provide the positive charge that stabilizes the oxyanion that is formed during catalysis. There is no analogous amino acid residue in the active site of the MMPs.

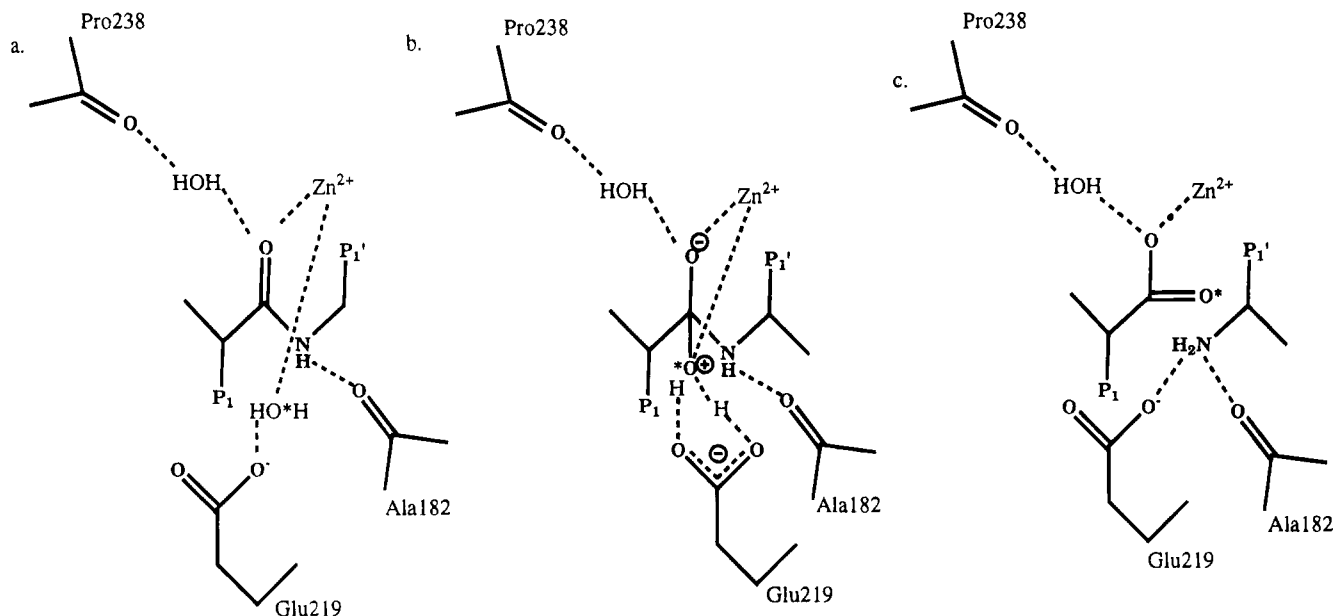


FIGURE 5: Proposed mechanism for matrilysin-catalyzed cleavage of a peptide substrate. (a) Dipeptide substrate is bound, and the presumed catalytic water (*) is shown interacting with both Glu219 and the active site zinc. The other water was identified in all structures of matrilysin with bound inhibitor and is hypothesized to stabilize the bound substrate. (b) Configuration of the tetrahedral intermediate. (c) Proposed interactions of matrilysin and the cleaved products.

Close examination of the matrilysin structures did not reveal an amino acid residue in the vicinity that could play such a role. In all three matrilysin–inhibitor complexes, however, there is a protein-bound water molecule which would be situated to form a hydrogen bond with the oxyanion of the tetrahedral intermediate (Table 3, H₂O₍₁₎ and Figure 5). When the active sites of matrilysin and thermolysin are superimposed, the conserved water molecule in matrilysin occupies approximately the same position as Ne of His231 of thermolysin. In similar enzyme–hydroxamate complexes, the Ne of His231 of thermolysin is 3.2 Å from the carbonyl oxygen of the hydroxamate (Holmes & Matthews, 1981), and the conserved water in matrilysin is 3.0 Å away (Figure 3B). There are also similarities to the position of this water molecule in matrilysin and Arg127 in carboxypeptidase A. For example, when there is a zinc-bound carboxylate in carboxypeptidase A, Arg127 is in position to make a hydrogen bond with the carboxylate (Christianson & Lipscomb, 1989). In an analogous manner, the carboxylate–matrilysin complex shows the water in position to interact with the zinc-bound carboxylate group (Figure 3B). The same structural similarity exists for the sulfodiimine complexes of matrilysin and carboxypeptidase A.

The absence of an amino acid residue in the active site of MMPs for the stabilization of the transition state makes this family of enzymes unique when compared to both thermolysin and carboxypeptidase A. It is possible that an unforeseen conformational change occurs in the MMPs during substrate binding or catalysis that repositions an amino acid residue such that it could participate in the stabilization of the transition state. There is, however, no evidence to suggest that this occurs. It is also possible that the absence of a positively charged residue to stabilize the transition state is favored in MMPs due to the unique zinc-coordinating environment (Lovejoy et al., 1994b). Three histidine residues coordinate the active site zinc in MMPs, whereas two histidines and one glutamic acid coordinate the zincs of thermolysin and carboxypeptidase A. The presence of a

conserved water in matrilysin underscores the role an additional group may have in stabilizing the transition state but suggests that this is perhaps not an essential feature of zinc-dependent proteases. Clearly, further investigation will be required to fully understand the catalytic mechanism of MMPs.

ACKNOWLEDGMENT

We thank Christine Luong and Rebecca Mackenzie for help with the crystallization experiments, Jim Barnett and co-workers for providing zymogen protein, and members of the Molecular Structure Department and the MMP project for valuable discussions and comments.

REFERENCES

- Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F., & Weng, J. (1987) in *Crystallographic Databases—Information Content, Software Systems, Scientific Applications* (Allen, F. H., Bergerhoff, G., & Sievers, R., Eds.) pp 107–132, Data Commission of the International Union of Crystallography, Bonn.
- Barnett, J., Straub, K., Nguyen, B., Chow, J., Suttman, R., Thompson, K., Tsing, S., Benton, P., Schatzman, R., Chen, M., & Chan, H. (1994) *Protein Expression Purif.* 5, 27–36.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol.* 112, 535–542.
- Birkedal-Hansen, H., Moore, W. G., Bodden, M. K., Windsor, L. J., Birkedal-Hansen, B., DeCarlo, A., & Engler, J. A. (1993) *Crit. Rev. Oral Biol. Med.* 4, 197–250.
- Blundell, T. L. (1994) *Struct. Biol.* 1, 73–75.
- Bode, W., Gomis-Ruth, F.-X., & Stockler, W. (1993) *FEBS Lett.* 331, 134–140.
- Bode, W., Reinemer, P., Huber, R., Kleine, T., Schnierer, S., & Tschesche, H. (1994) *EMBO J.* 13, 1263–1269.
- Borkakoti, N., Winkler, F. K., Williams, D. H., D'Arcy, A., Broadhurst, M. J., Brown, P. A., Johnson, W. H., & Murray, E. J. (1994) *Struct. Biol.* 1, 106–110.
- Brunger, A. T. (1987a) *Science* 235, 458–460.
- Brunger, A. T. (1987b) *X-PLOR: Version 3.1*, Yale University Press, New Haven, CT.
- Brunger, A. T. (1990) *Acta Crystallogr. A* 45, 46–57.

- Cappalonga, A. M., Alexander, R. S., & Christianson, D. W. (1992) *J. Biol. Chem.* 267, 19192–19197.
- CCP4 (1994) *Acta Crystallogr. D* 50, 760–763.
- Christianson, D. W. (1991) *Adv. Protein Chem.* 42, 281–355.
- Christianson, D. W., & Lipscomb, W. N. (1989) *Acc. Chem. Res.* 22, 62–69.
- Clark, I. M., & Cawston, T. E. (1989) *Biochem. J.* 263, 201–206.
- Fujinaga, M., & Read, R. J. (1987) *J. Appl. Crystallogr.* 20, 517–521.
- Gooley, P. R., O'Connell, J. F., Marcy, A. I., Cuca, G. C., Salowe, S. P., Bush, B. L., Hermes, J. D., Esser, C. K., Hagmann, W. K., Springer, J. P., & Johnson, B. A. (1994) *Struct. Biol.* 1, 111–118.
- Hodgkin, E. E., Gillman, I. C., & Gilbert, R. J. (1994) *Protein Sci.* 3, 984–986.
- Holmes, M. A., & Matthews, B. W. (1981) *Biochemistry* 20, 6912–6920.
- Housely, T. J., Baumann, A. P., Braun, I. D., Davis, G., Seperack, P. K., & Wilhelm, S. M. (1993) *J. Biol. Chem.* 268, 4481–4487.
- Jancarik, J., & Kim, S.-H. (1991) *J. Appl. Crystallogr.* 24, 409–411.
- Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983) *Science* 220, 671–680.
- Knight, C. G., Willenbrock, F., & Murphy, G. (1992) *FEBS Lett.* 296, 263–266.
- Lovejoy, B., Cleasby, A., Hassell, A. M., Longley, K., Luther, M. A., Weigl, D., McGeehan, G., McElroy, A. B., Drewry, D., Lambert, M. H., & Jordan, S. R. (1994a) *Science* 263, 375–377.
- Lovejoy, B., Hassell, A. M., Luther, M. A., Weigl, D., & Jordan, S. R. (1994b) *Biochemistry* 33, 8207–8217.
- Lowry, C. L., McGeehan, G., & LeVine, H. (1992) *Proteins* 1, 42–48.
- Luzzati, V. (1952) *Acta Crystallogr.* 5, 802–810.
- Matrisian, L. M. (1992) *BioEssays* 14, 455–463.
- Matthews, B. W. (1988) *Acc. Chem. Res.* 21, 333–340.
- McGrath, M. E., Haymore, B. L., Summers, N. L., Craik, C. S., & Fletterick, R. J. (1993) *Biochemistry* 32, 1914–1919.
- McPherson, A. (1982) *Preparation and analysis of protein crystals*, John Wiley & Sons, New York.
- Monzingo, A. F., & Matthews, B. W. (1984) *Biochemistry* 23, 5724–5729.
- Murphy, G., Allan, J. A., Willenbrock, F., Cockett, M. I., O'Connell, J. P., & Docherty, A. J. (1992) *J. Biol. Chem.* 267, 9612–9618.
- Netzel-Arnett, S., Fields, G. B., Birkedal-Hansen, H., & VanWart, H. E. (1991) *J. Biol. Chem.* 266, 6747–6755.
- Netzel-Arnett, S., Sang, Q.-X., Moore, W. G. I., Navre, M., Birkedal-Hansen, H., & VanWart, H. E. (1993) *Biochemistry* 32, 6427–6432.
- Okada, Y., Nagase, H., & Harris, E. D. (1986) *J. Biol. Chem.* 261, 142245–142255.
- Phillips, M. A., Fletterick, R., & Rutter, W. J. (1990) *J. Biol. Chem.* 265, 20692–20698.
- Reinemer, P., Grams, F., Huber, R., Kleine, T., Schnierer, S., Piper, M., Tschesche, H., & Bode, W. (1994) *FEBS Lett.* 338, 227–233.
- Rossmann, M. G. (1990) *Acta Crystallogr. A* 46, 73–82.
- Sack, J. S. (1988) *J. Mol. Graphics* 6, 244–245.
- Sanchez-Lopez, R., Alexander, C. M., Behrendtsen, O., Breathnach, R., & Werb, Z. (1993) *J. Biol. Chem.* 268, 7238–7247.
- Springman, E. B., Angleton, E. L., Birkedal-Hanssen, H., & VanWart, H. E. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 364–368.
- Spurlino, J. C., Smallwood, A. M., Carlton, D. D., Banks, T. M., Vavra, K. J., Johnson, J. S., Cook, E. R., Falvo, J., Wahl, R. C., Pulvino, T. A., Wendoloski, J. H., & Smith, D. L. (1994) *Proteins* 19, 98–109.
- Stams, T., Spurlino, J. C., Smith, D. L., Wahl, R. C., Ho, T. F., Qoronfleh, M. W., Banks, T. M., & Rubin, B. (1994) *Struct. Biol.* 1, 119–123.
- Stetler-Stevenson, W. G., Aznavoorian, S., & Liotta, L. A. (1993) *Annu. Rev. Cell Biol.* 9, 541–573.
- VanWart, H. E., & Birkedal-Hansen, H. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 5578–5582.
- Windsor, L. J., Birkedal-Hansen, H., & Engler, J. A. (1991) *Biochemistry* 30, 641–647.

BI9425819