PROTEINS: Structure, Function, and Genetics 19:98–109 (1994)

1.56 Å Structure of Mature Truncated Human Fibroblast Collagenase

John C. Spurlino,[1] Angela M. Smallwood,[1] Dennis D. Carlton,[2] Tracey M. Banks,[1] Karen J. Vavra,[2] Jeffrey S. Johnson,[1] Ewell R. Cook,[1] Joseph Falvo,[2] Robert C. Wahl,[1] Tricia A. Pulvino,[1] John J. Wendoloski,[1] and Douglas L. Smith[1]

[1]Sterling Winthrop Pharmaceuticals Research Division, Collegeville, Pennsylvania 19426, and [2]Eastman Kodak Company, Rochester, New York 14650

ABSTRACT **The X-ray crystal structure of a 19 kDa active fragment of human fibroblast collagenase has been determined by the multiple isomorphous replacement method and refined at 1.56 Å resolution to an *R*-factor of 17.4%. The current structure includes a bound hydroxamate inhibitor, 88 waters and three metal atoms (two zincs and a calcium). The overall topology of the enzyme, comprised of a five stranded β-sheet and three α-helices, is similar to the thermolysin-like metalloproteinases. There are some important differences between the collagenase and thermolysin families of enzymes. The active site zinc ligands are all histidines (His-218, His-222, and His-228). The presence of a second zinc ion in a structural role is a unique feature of the matrix metalloproteinases. The binding properties of the active site cleft are more dependent on the main chain conformation of the enzyme (and substrate) compared with thermolysin. A mechanism of action for peptide cleavage similar to that of thermolysin is proposed for fibroblast collagenase.** © 1994 Wiley-Liss, Inc.

Key words: crystallography, hydroxamate, high resolution, metalloproteinase, zinc, X-ray

INTRODUCTION

Collagenases are members of the matrix metalloproteinase (MMP) family of enzymes (for a recent review see Woessner's paper[1]). Members of the MMP family are distinguished by several features: a highly similar N-terminal catalytic domain, requirement for Ca^{2+}, a zinc atom at the active site, and secretion as inactive proenzymes which are activated by the cleavage of approximately 80 amino acids from the N-terminus.[2]

MMP enzymes have been implicated in a variety of disorders involving the proteolytic breakdown of connective tissue.[3] These range from periodontal disease and arthritis to cancer tumor metastasis. Fibroblast collagenase has the ability to cleave the triple helical types I, II, and III interstitial collagens at a specific site.[1,4] This ability marks fibroblast collagenase as a prime target for control of arthritis.

The domain structure of human fibroblast collagenase consists of a 19 residue signal peptide, a propeptide of 80 residues, the catalytic domain (residues 100–269), and the hemopexin-like domain (residues 270–469). The numbering of residues is based on the sequence of Whitham et al.[6] Activation of the enzyme produces N-terminal residues at positions 100–102, with the highest specific activity observed for the N-terminal residue of Phe-100.[1] The full length (residues 100–469) active unglycosolated human fibroblast collagenase (hfc), 42 kDa, is unstable, presumably due to autolysis. This instability make the full length mature hfc a difficult target for crystallographic studies. However, a C-terminally truncated form of hfc produced via recombinant DNA technology[5] forms a stable activated 19 kDa enzyme (residues 101–269 of the native enzyme[6]) well suited for X-ray studies. This mature (activated) truncated collagenase (mCL-t) has a specific activity and K_m for peptides comparable to the native enzyme. However, mCL-t does not cleave type I collagen. The construct of Lowry[7] appears to produce an identical enzyme to the mCL-t used here.

The refined, high resolution structure of mCL-t reveals the active site architecture and serves as a basis for the understanding of mechanism of proteolysis of collagen and its dependence on Ca^{2+} and Zn^{2+} ions.

MATERIALS AND METHODS

Materials

Ammonium sulfate was from Schwartz Mann. The hydroxamate inhibitor (see Fig. 1) and the *p*-iodo derivative of it were synthesized by Ewell Cook and Jeff Johnson, respectively (Sterling Winthrop Phar-

Received December 6, 1993; revision accepted January 3, 1994.

Address reprint requests to John C. Spurlino, Sterling Winthrop Pharmaceuticals Research Division, 1250 S. Collegeville Rd., P.O.Box 5000, Collegeville, PA 19426-0900.

maceutical Research Division, Medicinal Chemistry Division). All other reagents and chemicals were of analytical reagent grade.

Protein Purification and Crystallization

Purified protein was obtained from a 10 liter fermentation of *E. coli* transformed with pKV150.[5] Procollagenase (C-terminally truncated at Pro-269) with the addition of an N-terminal methionine was obtained in inclusion bodies. The inclusion bodies were isolated, urea solubilized, and purified by anion exchange chromatography (monoQ 16/10 column, Pharmacia). The partially purified protein was refolded, activated by aminophenylmercuric acetate incubation, and further purified by gel filtration chromatography (S100HR resin, Pharmacia).

Crystallization was accomplished by the batch method. A protein solution containing 15 mg/ml mCL-t (ϵ = 1.34 mg/cm) in 2.5 mM Tris, 5.0 mM $CaCl_2$, 0.4 M NaCl, 3 mM NaN_3, 1 mM hydroxamate inhibitor (see Fig. 1), pH 7.5 was prepared. A small vial containing 100 µl of this protein solution was titrated with solid ammonium sulfate until a precipitate formed. The resulting solution was back titrated with distilled water until the precipitate redissolved. The vial was capped and clusters of crystals formed within a week. The crystals of mCL-t have space group 19 ($P2_12_12_1$), with cell dimensions of a = 109.21 Å, b = 44.57 Å, c = 36.28 Å. The crystals used for data collection had average dimensions of 0.2 × 0.4 × 0.8 mm and a rhomboid cross section. These crystals diffract to at least 1.5 Å resolution.

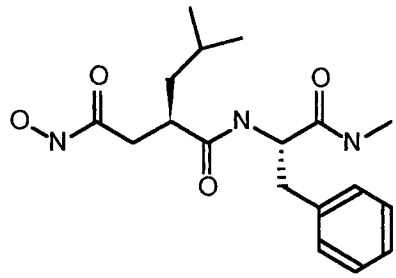

Fig. 1. The structure of the hydroxamate inhibitor used in the crystallization of mCL-t.

Structural Methods

Diffraction data were collected on a Siemens 512 × 512 multiwire area detector (CuK_α radiation, Rigaku RU200 rotating anode generator, 50 kV, 100 mA, 0.3 × 3 mm focal spot, double focusing mirrors) and processed with XENGEN.[8,9] Diffraction data were also collected on a MAR imageplate area detector (CuK_α radiation, Rigaku RU200 rotating anode generator, 50 kV, 100 mA, 0.3 × 3 mm focal spot, graphite monochromator) and processed with MOSFLM.[10] All data were collected at room temperature. There was minor radiation decay with average intensities decreasing by 30% after 6 days of X-ray exposure. Data merging, heavy atom substi-

TABLE I. mCL-t Crystal Data Statistics (MAR detector)

Data (2.6 Å)	Conc./time (mM/days)	Unit cell (Å)			Sites	F_H / residual
		a	b	c		
Native		109.21	44.57	36.28		
DyI_3	2.0/14.0	109.16	44.46	36.40	2	1.88
$Sm_2(SO_4)_3$	2.0/16.9	109.04	44.45	36.28	2	2.21
Iodinated inhibitor	sat./1.0	109.36	44.55	36.27	1	1.15
KI/I_3	*/14.0	109.57	44.46	36.54	2	1.56

*Direct iodination of protein tyrosines.[15]

TABLE II. Data Statistics for Native mCL-t (Siemens)

Lower limit shell (Å)	Number of measurements	Number of reflections	Percentage of possible data	R_{sym}*
2.83	16,784	4,641	99.1	4.24
2.25	16,737	4,322	97.1	6.30
1.97	9,390	4,146	93.4	7.76
1.79	8,211	3,897	89.1	11.97
1.66	7,349	3,688	84.8	19.35
1.56	5,791	3,218	73.7	26.88
Total	64,262	23,912	89.7	6.42

*$R_{Sym} = \sum_r^N \left(\sum_i^{(r)} |I(i,r)| - \langle I(r) \rangle \right) / \sum_r^N (n(r) \times \langle I(r) \rangle) \times 100$ where N is the number of unique reflections, $n(r)$ is the number of multiple measurements for the rth reflection with the mean value $\langle I(r) \rangle$, and $I(i,r)$ is the ith individual measurement of the rth reflection.

TABLE III. mCL-t Refinement

	R-factor*	Residues	Cofactors	Waters
Initial MIR fit (2.6 Å resolution)	0.512	133	None	0
XPLOR slow cool refinement (2.6 Å resolution)	0.319	133	None	0
MIR + $2F_o-F_c$ map Refit XPLOR slow cool refinement (2.6 Å resolution)	0.301	159	2 Zn + 1 Ca	0
$2F_o-F_c$ map Refit PROLSQ refinement (2.0 Å resolution)	0.268	164	2 Zn + 1 Ca	0
$2F_o-F_c$ map Refit PROLSQ refinement (1.56 Å resolution)	0.200	164	2 Zn + 1 Ca and inhibitor	0
$2F_o-F_c$ map Refit PROLSQ refinement (1.56 Å resolution)	0.174	160	2 Zn + 1 Ca and inhibitor	88

$$*R-\text{factor} = \frac{\Sigma||F_{obs}| - |F_{calc}||}{\Sigma|F_{obs}|}.$$

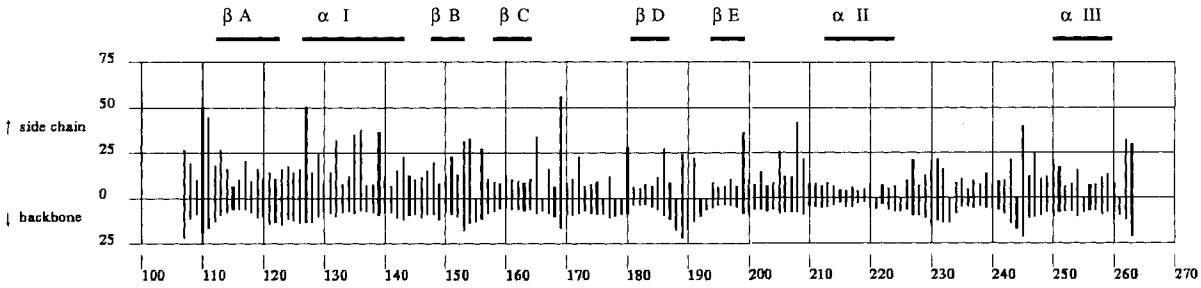

Fig. 2. B-value plot for the refined mCL-t structure with the secondary structure marked above.

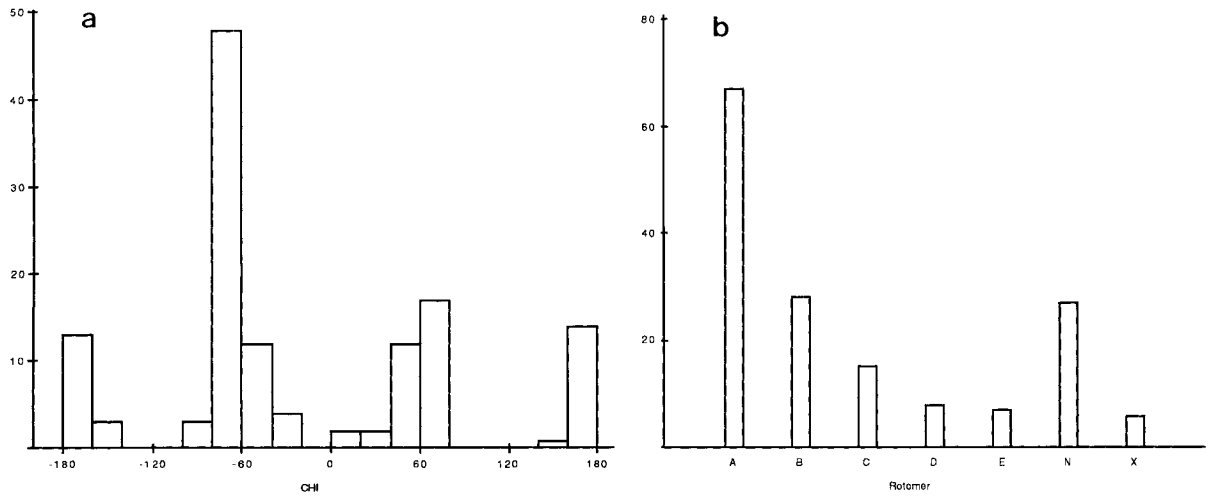

Fig. 3. (**a**) CHI_1 distribution for mCL-t. (**b**) Rotamer[23] distribution for mCL-t. A-E are the five most common occurrences of the rotamer for the appropriate residue type, N is when no rotamer designation is available (Gly and Ala residues), and X is for the sixth or higher rotamer conformation or a nonrotamer conformation.

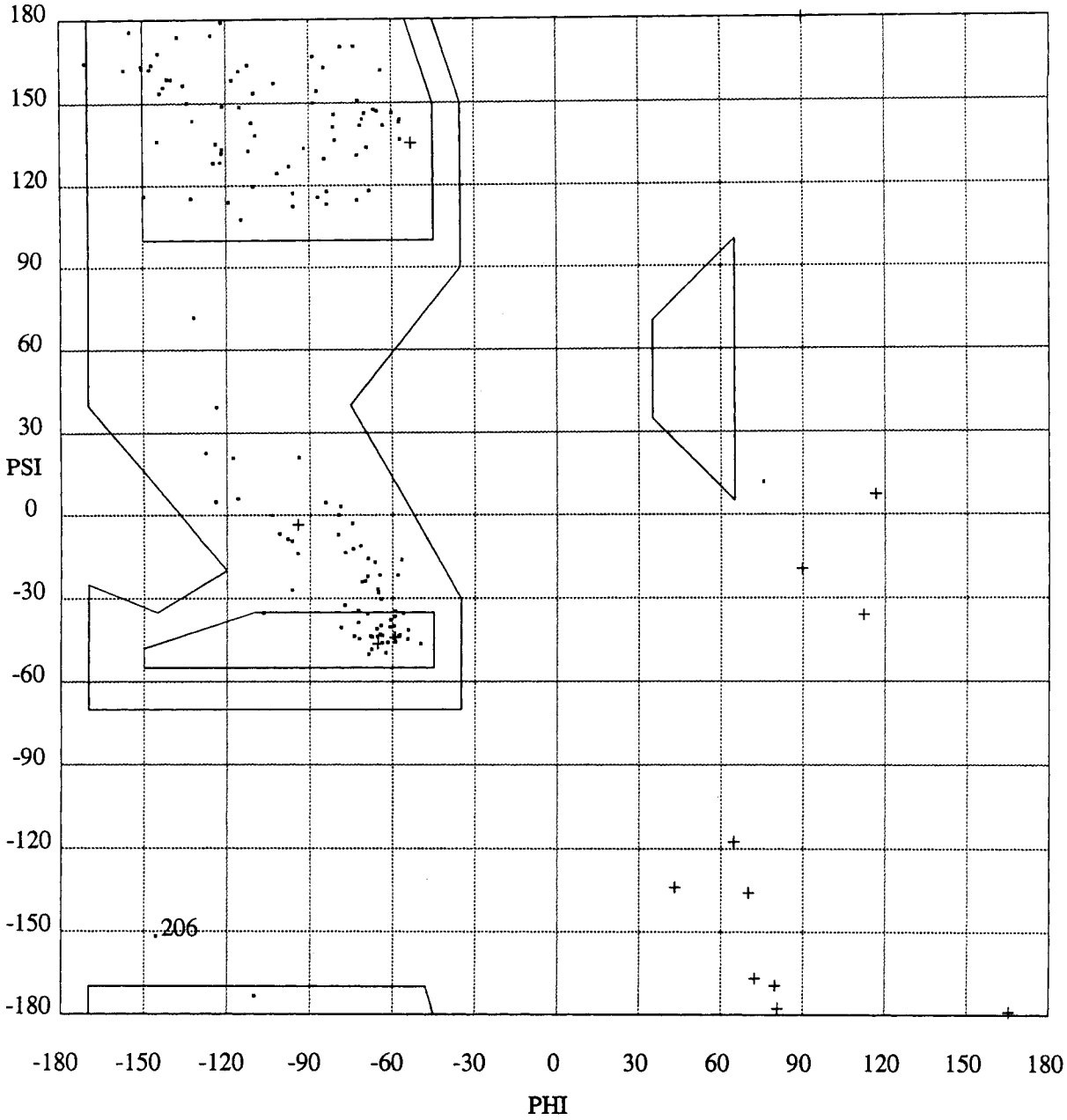

Fig. 4. Ramachandran plot for mCL-t. Glycine residues are marked as + and all others by a dot. Residue 206 (labeled) is the only outlier.

tution and MIR calculations were performed using PROTEIN 3.[11] Refinement of the protein structure was initiated using the XPLOR software package[12] and completed with PROLSQ.[13] All fitting was performed with the crystallographic software package CHAIN.[14] The initial protein trace was fitted to an F_{obs},α_{MIR} map. Intermediate and final fitting were done with $(2|F_{obs}| - |F_{calc}|)$ and $(|F_{obs}| - |F_{calc}|)$ coefficients and α_{calc} phases from restrained least-squares refinement.

RESULTS AND DISCUSSION

Refinement

MIR analysis was conducted at 2.6 Å resolution based on 4 heavy atom derivatives (see Table I). The mean figure of merit ($<m>$) of the 2.6 Å MIR analysis was 0.68. The dysprosium and samarium sites were coincidental, but, since the occupancies were different between the sites in the two derivatives, they were both used in phasing. Direct iodination of

TABLE IV. Refinement Statistics for mCL-t

Resolution	10.0–1.56Å
$R = \frac{\Sigma\|\|F_{obs}\| - \|F_{calc}\|\|}{\Sigma\|F_{obs}\|}$	17.4
Average $\|\|F_o\|-\|F_c\|\|$	22.11
No. unique reflections	21,710
No. of nonhydrogen atoms	1364

rms deviation in	mCL-t	Target σ
Bond distances (Å)	0.020	0.020
Angle distances (Å)	0.049	0.050
Planar 1–4 distances (Å)	0.055	0.050
Planarity (Å)	0.012	0.015
Chiral volume (Å^3)	0.221	0.200
Torsion angles (deg)		
Planar (0, 180)	2.5	3.0
Staggered (±60, 180)	14.0	15.0
Orthonormal (±90)	26.9	20.0
Nonbonded contacts (Å)		
Single torsion	0.176	0.350
Multiple torsion	0.158	0.350
H-bond	0.141	0.350
Thermal parameters (Å^2)		
Main chain bond	2.126	2.500
Main chain angle	2.906	4.500
Side chain bond	4.582	4.000
Side chain angle	6.819	6.000

tyrosines in the enzyme[15] was used to obtain one derivative. The iodinated inhibitor identified in Table I is the para-iodo derivative of the compound in Figure 1.

The initial model from the MIR map consisted of a polyalanine chain of 133 residues. Subsequently side chains were fitted to electron density where unambiguous interpretation was possible. This model was subjected to simulated annealing refinement resulting in an R-factor of 0.301. Several rounds of simulated annealing refinement[12] at 2.6 Å resolution followed by refitting were required to essentially complete the chain trace.

At this point a complete 1.56 Å data set, obtained from a single crystal, (see Table II) was used to continue refinement. A stepped extension of the resolution from 2.6 Å to 1.56 Å was performed with the restrained least-squares program PROLSQ.[13] The first six and the last six residues (101–106 and 264–269) are not clearly evident in the electron density maps and therefore were omitted from the structure in the final refinement. A synopsis of the major steps in refinement is shown in Table III.

Water molecules were added where 2σ density for both F_o-F_c and $2F_o-F_c$ maps was evident within 3.5 Å of a suitable protein atom. Water molecules were deleted from the refinement if their B value exceeded 50 Å^2. The geometry of the refined structure can be judged from Table IV. Other indicators of the quality of the structure are a B-value plot (Fig. 2), Chi and rotamer histograms (Fig. 3), and a Ramachandran plot (Fig. 4).

Cis-peptides are infrequently reported in protein structures. Out of 10,813 ω angles reported in structures between 2.0 and 1.5 Å resolution only 0.38% were cis and of those 88% involved prolines.[16] Near the completion of refinement the ω torsion angle for the Glu-209–Tyr-210 peptide bond was noticed to deviate greatly from 180°. Changing the constraints to a cis-peptide for the bond between Glu-209 and Tyr-210 resulted in a smooth refinement to ω =

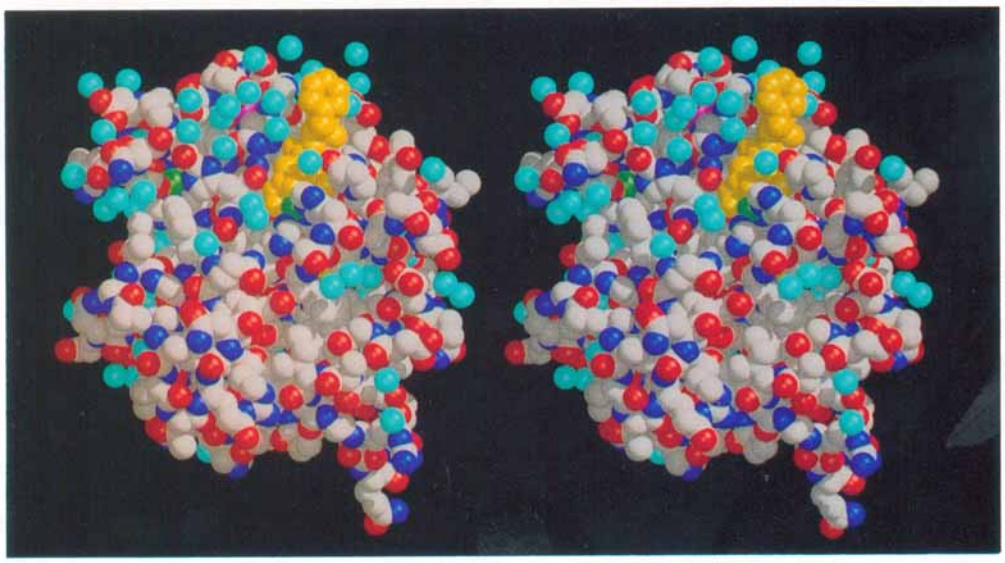

Fig. 5. Stereo CPK rendering of mCL-t. Water molecules are cyan colored, zinc ions are green, the calcium ion is magenta, the hydroxamate inhibitor is colored yellow. The remaining atoms are colored by atom type; carbon is gray, nitrogen is blue, oxygen is red, and sulfur is dark yellow.

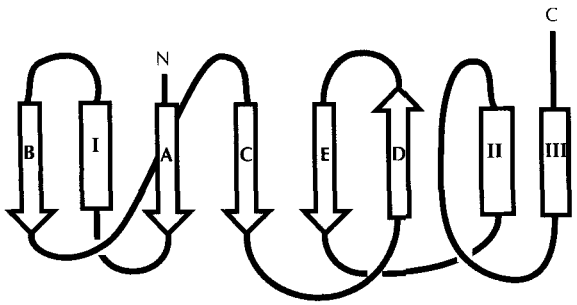

Fig. 6. A topological drawing of mCL-t. The strands are depicted as arrows and labeled with capital letters and the helices are shown as rectangles labeled with Roman numerals.

−1.0°. The presence of this cis peptide allows one additional hydrogen bond between the carbonyl oxygen of Arg-202 and the amide nitrogen of Tyr-210 that would not be possible otherwise. This bond is not near the active site zinc, although residues 209 and 210 are located near the probable P3′ pocket.

Structure

The structure of mCL-t is an oblate sphere with one creased face and dimensions of 35 × 30 × 30 Å. The active site is located centrally within a 20 Å groove which runs along the face of the sphere. A stereo CPK rendering of mCL-t (Fig. 5) shows the active site groove and the orientation of the hydroxamate inhibitor. The enzyme folds as a five-stranded (mostly parallel) β-sheet flanked on one side by three α-helices (see Figs. 6 and 7). The active site groove is formed between the antiparallel strand D and helix II. The loop connecting helix II and helix III also serves to define the binding groove.

The loop between strands C and D consists of three turns which contain residues vital to the coordination of one zinc ion and the calcium ion. The first turn positions the structural zinc binding loop at the top of the enzyme, bringing His-168 and Asp-170 into position to ligand a zinc ion. The third turn forms the calcium binding loop and contains three of the six residues involved in liganding the calcium ion. The first residue of strand D also forms a bond with the calcium ion.

The N-terminus of mCL-t is located in close proximity to the C-terminus and the active cleft. The position of the cleaved N-terminus is consistent with the switching mechanism in which a conserved cysteine (residue 95) forms the fourth zinc ligand in the inactive pro-form of the enzyme.[18] The C-terminus of the truncated enzyme offers little information as to the possible location of the hemopexin-like domain in the full length enzyme.

Several neutral proteinases (thermolysin, elastase, and neutral protease) have structural similarity.[19] Fibroblast collagenase is another neutral protease and has many of the structural features seen in the thermolysin family. The astacin family of endopeptidases are closer in structural arrangement and active site zinc ligands to the collagenases than thermolysin. However, as the coordinates of the astacin structures are not currently available a precise comparison with astacin is not possible, therefore the best model for a structural comparison is thermolysin. The basic fold of the active site follows the same topology (strand C–helix II). The basic arrangement of secondary structure, an N-terminal domain of mixed β-sheet with an α-helical C-domain, is also the same (see Fig. 8).

The comparison of α-carbon atom coordinates of thermolysin (pdb7tln)[20] with mclt using OVRLAP[21] showed an rms deviation of 1.34 Å for 32 residues (see Fig. 9). The extension of the structural comparison produced an alignment of 72 residues with an rms deviation of 2.8 Å. The broader comparison covers all secondary structural elements except strand B and includes a portion of the loop between helices II and III. The similarity of topological arrangement is most prominent with the N-domain of thermolysin. Although there is an alignment with helix III, the C-domain of thermolysin is much larger than that of mCL-t and the structural arrangement that defines that side of the active site is different. There are several other notable differences between collagenase and the thermolysin-like neutral proteases. These differences primarily involve metal coordination and inhibitor binding.

Zinc Coordination

The number of zinc ions in collagenases has been debated. The structures of astacin[22] and adamalysin II[23] show only a single zinc ion. The work of Lowry[7] indicated only a single zinc ion in mCL-t. In contrast the stoichiometric measurements of Salowe et al.[24] indicated two zinc ions in stromelysin. The structural work presented here clearly shows that there are two zinc ions in mCL-t. The high sequence homology between the collagenases (and stromelysin) indicates that two zinc ions will be found in all their structures. The presence of a second zinc atom differentiates the collagenases from thermolysin and astacin families of enzymes. In addition the ligands to the active site zinc are different from those observed in thermolysin. The zinc atoms will be referred to as Zn1 (the zinc which is coordinated to the bound inhibitor) and Zn2. The two zinc ions are 12.3 Å apart.

The conserved MMP sequence HExGHxxGxxH[1] contains the ligands for the active site zinc in collagenase. Zn1 is coordinated by the imidazole nitrogen atoms of His-218, His-222, and His-228 and the two hydroxamate oxygens of the bound inhibitor with an average metal–ligand distance of 2.09 Å (see Table V). The geometry of the coordination can best be described as square pyramidal (or pentacoordinate) with the two hydroxamate oxygen atoms occupying

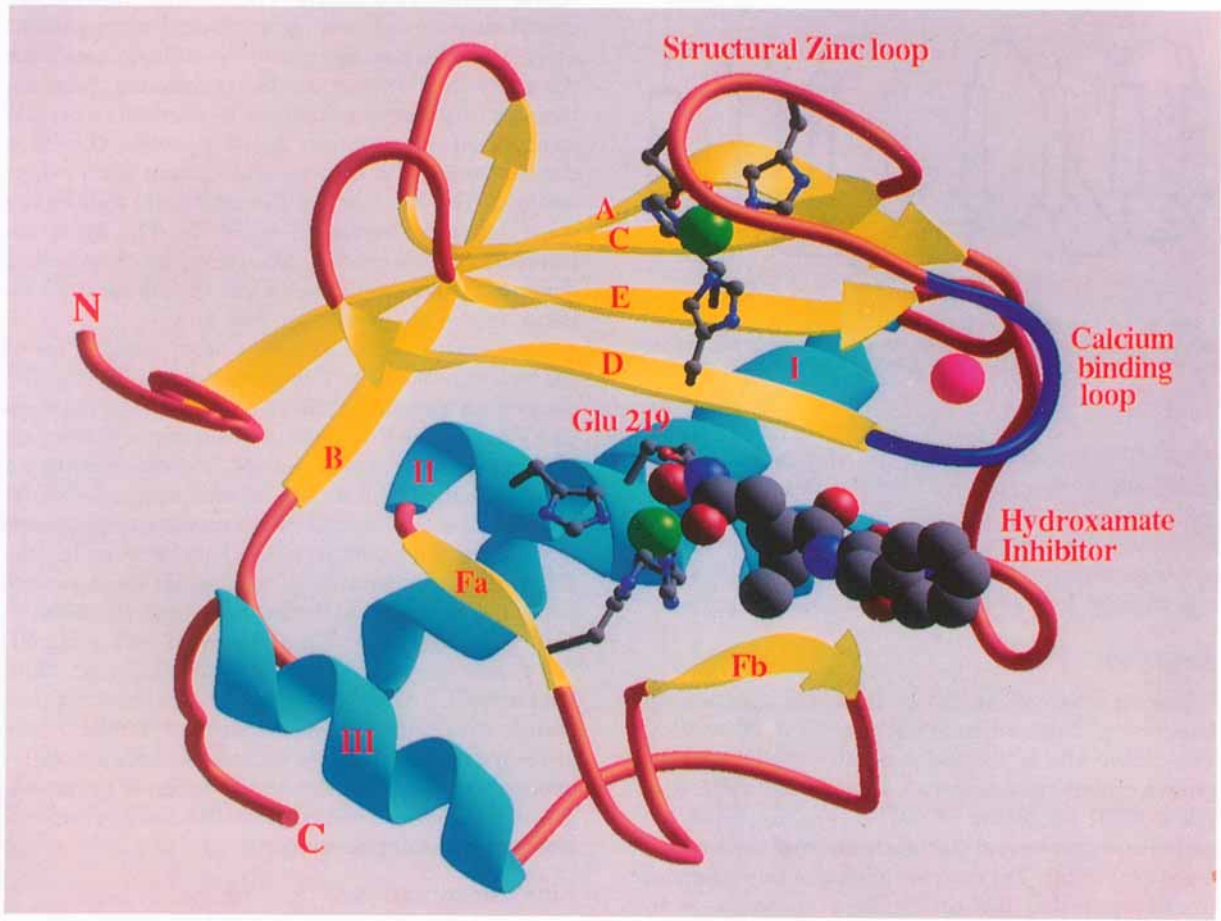

Fig. 7. Ribbon Drawing of mCL-t.[17] β-Strands are labeled with capital letters and α-helices are labeled with Roman numerals in the order of appearance.

two of the base coordination points. This pentacoordinate geometry was first described for thermolysin in conjunction with hydroxamate inhibitors[25] as indicative of an intermediate in peptide hydrolysis.

The second glycine of the conserved active-site sequence occurs at the C-terminus of the helix containing the first two histidines and active glutamate residue which allows the chain to fold back into the active site with the third histidine to complete the zinc binding triad. The spacing between the third and second histidine is the same as reported for the astacin structure.[22] The three histidines of mCL-t are oriented similarly to those of astacin, however astacin was solved in an uninhibited form and displayed tetrahedral coordination (with a water molecule) at the active site zinc.

The positions of two of the three histidines (218 and 222) are further stabilized through hydrogen bonds formed with other residues in the protein. Histidine 218 forms a hydrogen bond with the carbonyl oxygen of Leu-235 (C–O→ND1 angle = 142°, distance = 2.8 Å). The chi angles for His-218 match Ponder and Richards[26] rotamer #5. Histidine 222 also forms a hydrogen bond between the ND1 nitrogen and a carbonyl oxygen atom of Leu-226 (C–O→ND1 angle = 142°, distance = 2.9 Å). The chi angles for His-222 are closest to rotamer #1, although chi2 differs by 20° (ς = 17.2) from the average value for rotamer #1. Histidine 228 which has the protonated ND1 located at the protein surface has the greatest deviation from a known rotamer. Chi2 deviates from rotamer #3 by 46° (σ = 32.2).

Zn2 is tetrahedrally coordinated by the imidazole nitrogens of His-168, His-183, and His-196 and one of the carboxyl oxygens of Asp-170 (see Table VI). The average metal–ligand distance is 2.02 Å (metal–histidine ligand distance = 2.07 Å). The coordination and inaccessibility of the surface of Zn2 suggest that its role is structural rather than catalytic. This second zinc ion appears to be a feature that is unique to the collagenase family of metalloproteinases.

Histidine 168 forms a hydrogen bond between the ND1 atom and the carbonyl oxygen of Gly-166. Likewise the other oxygen of Asp-170 is involved in a hydrogen bond with the backbone nitrogen of Ser-

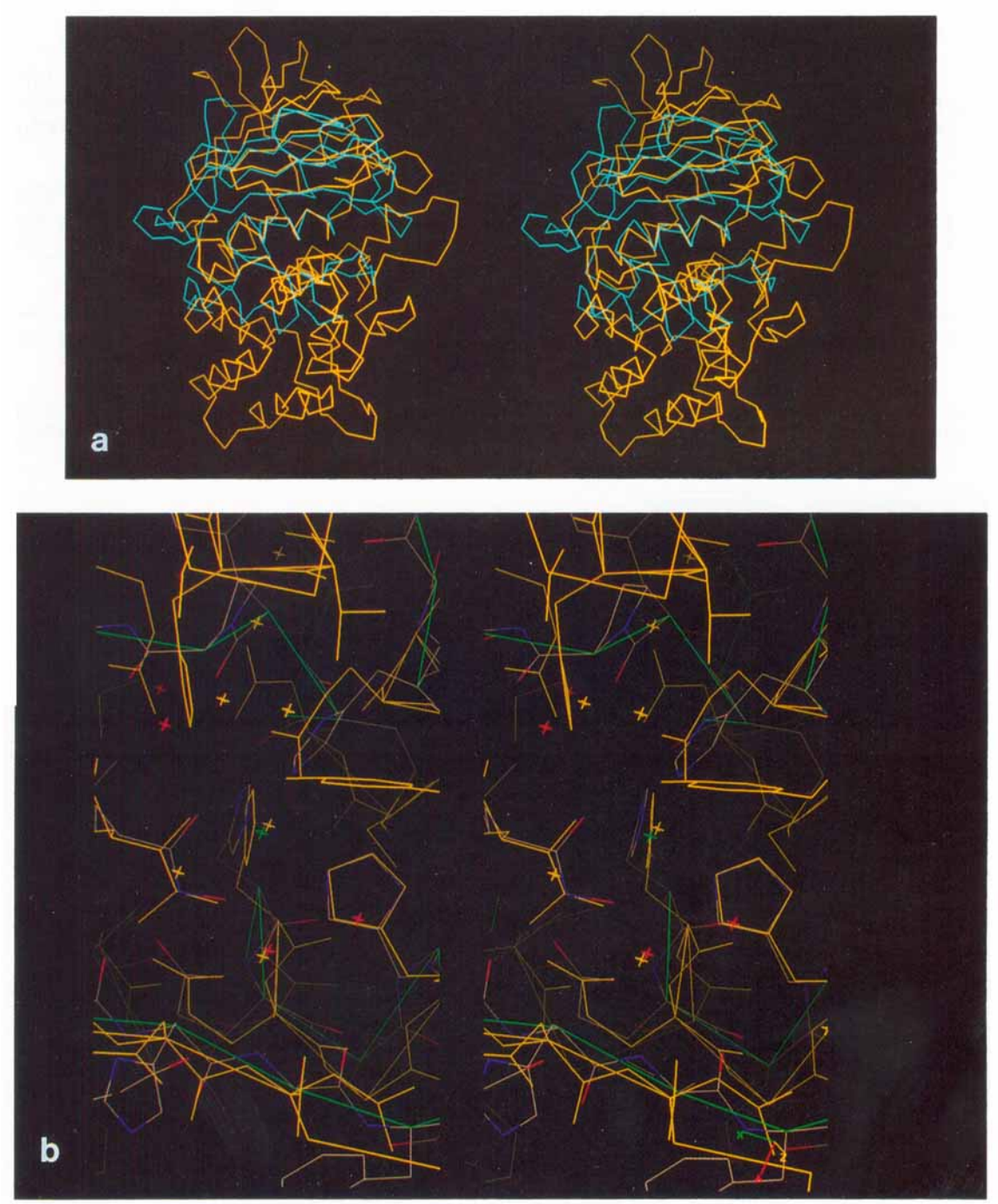

Fig. 8. (**a**) Superimposition of the α-carbon traces of mCL-t (blue) and thermolysin (yellow) calculated with the program OVRLAP.[19] (**b**) The active site regions of mCL-t and thermolysin (yellow) using the same superimposition as in **a.**

172. These hydrogen bonds help to stabilize the conformation of the respective residue for optimum interaction with Zn2. The remaining histidine residues involved in ligand interactions with Zn2 are also further stabilized, but by hydrogen bonds achieved indirectly through water.

Calcium Binding

Only one calcium ion is found in the truncated enzyme. The calcium ion, which has been proven to be essential for enzymatic activity,[7] is octahedrally coordinated (see Table VII). The calcium atom stabilizes the glycine-rich loop (residues 175–180)

```
TLN     1 sequence   ITGTSTVGVG RGVLGDQKNI NTT

          structure              EEEEEEE  TT                          HHHHHHHHHHHHHH HHTT
MCLT  106 sequence       NPRWEQTHLTYRIENYTPDLP--------------------RADVDHAIEKAFQL-WSNVTP
TLN    24 sequence   YSTYYYLQDNTRGDGIFTYDAKYRTTLPGSLWADADNQFFASYDAPAVDAHYYAGVTYDYYKNVHNR

          structure   EEEEE     EEEEEEE                    EEEEE    TTTT EEEETT      EEEEE
MCLT  147 sequence   LTFTKVSEGQADIMISFVRGDHRDNSPFDGPGGNLAHAFQPGPGIGGDAHFDED----ERWTNNF
TLN    91 sequence   LSYDGNNA---AIRSSVHYSQG----------YNNAFWNG------SEMVYGDGDGQTFIPLS--

          structure     HHHHHHHHHHHHHT
MCLT  208 sequence   REYNLHRVAAHELGHSLGL----------SHST----------------------------
TLN   135 sequence   ---GGIDVVAHELTHAVTDYTAGLIYQNESGAINEAISDIFGTLVEFYANKNPDWEIGEDVYTPGI

          structure                                        TT TT
MCLT  231 sequence   ----------------------------------------DIGALMYPSY---------TFSGDVQL
TLN   198 sequence   SGDSLRSMSDPAKYGDPDHYSKRYTGTQDNGGVHINSGIINKAAYLISQGGTHYGVSVVGIGRDKL

          structure   HHHHHHHHHHH
MCLT  249 sequence   AQDDIDGIQAIYGRS
TLN   264 sequence   GKIFYRALTQYLTPTSNFSQLRAAAVQSATDLYGSTSQEVASVKQAFDAVGVK
```

Fig. 9. Topological alignment of mCL-t and thermolysin (TLN). Residues in bold indicate the tight comparison (rms = 1.34 Å) the boxed residues are those included in the extended comparison (rms = 2.8 Å). The secondary structure of mCL-t is indicated above (E = β-sheet, H = α-helix, and T = turn).

TABLE V. Zn1 Ligands and Their Spacings

Residue	Atom	*B*-value (Å^2)	M . . . L* (Å)	Ligand–metal–ligand angles (degree)			
				His-222	His-228	Inhib O1	Inhib O2
His-218	NE2	2.13	2.00	100.0	105.9	100.4	103.8
His-222	NE2	6.24	2.13		95.7	86.1	152.7
His-228	NE2	3.51	2.07			152.9	90.5
Inhibitor	O1	5.80	2.17				76.7
Inhibitor	O2	4.12	2.08				

*Metal–ligand distance.

TABLE VI. Zn2 Ligands and Their Spacings

Residue	Atom	*B*-value (Å^2)	M . . . L* (Å)	Ligand–metal–ligand angles (degree)		
				His-183	His-196	Asp-170
His-168	NE2	5.23	2.09	114.1	108.4	104.7
His-183	NE2	6.11	2.03		112.5	120.4
His-196	ND1	5.24	2.08			94.6
Asp-170	OD1	7.73	1.88			

*Metal–ligand distance.

TABLE VII. Calcium Ligands and Their Spacings

Residue	Atom	*B*-value (Å^2)	M . . . L* (Å)	Ligand–metal–ligand angles (degree)				
				Gly-176	Gly-178	Asn-180	Asp-198	Glu-201
Asp-175	OD1	8.77	2.30	88.4	86.7	87.2	93.5	176.3
Gly-176	O	9.01	2.33		84.6	174.6	88.1	91.9
Gly-178	O	9.78	2.44			92.0	172.7	89.6
Asn-180	O	7.57	2.39				95.3	92.3
Asp-198	OD2	7.83	2.32					90.2
Glu-201	OE2	8.29	2.22					

*Metal–ligand distance.

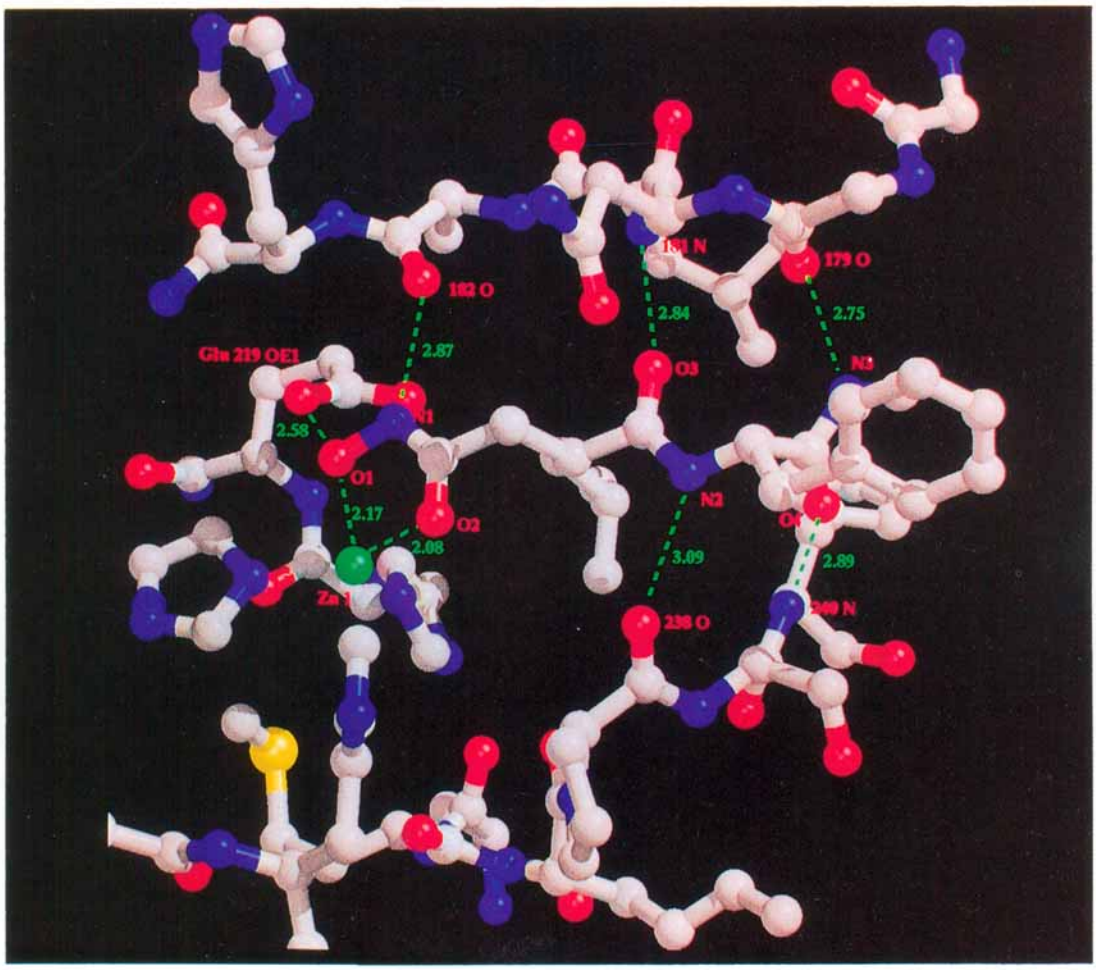

Fig. 10. Hydrogen bonding pattern for the hydroxamate inhibitor bound in mCL-t.

which anchors the portion of the β-strand D (179–182) involved in inhibitor binding.

The major site for both rare earth derivatives was coincident with the position of the calcium atom. There are no additional calcium atoms present in the mCL-t structure. Although there are four stabilizing calcium ions present in thermolysin, none of them are in an equivalent position to the single calcium found in mCL-t. The lone calcium ion reported for astacin[22] and adamalysin II[23] is also located in a structural position that is disparate from mCL-t.

Inhibitor Binding

The hydroxamate found at the active site not only makes two bonds with Zn1, but is further stabilized by six hydrogen bonds with the enzyme (see Fig. 10). Indeed, the peptide-like portion of the inhibitor can be envisioned as an extension of the β-sheet formed by strands A–E. Residues 179, 181, and 182 from strand D are H-bonded to the inhibitor. However, the active site cleft extends past the zinc along strand D and further H-bonding may play a role in the cleavage of collagen. Residues 238 and 240 also form hydrogen bonds with inhibitor. These residues are found on the opposite side of the active cleft from strand D and can be portrayed as the top strand of an extended β-sheet that includes the substrate (or inhibitor). This portion of the structure has been labeled as Fb (see Fig. 7) and the portion of structure that could possibly play the same role for the P-side of the peptide substrate is labeled as Fa.

Another difference between thermolysin and mCL-t involves the hydrogen bonding between the inhibitor and enzyme. Thermolysin utilizes side chains of residues lining the active cleft,[25,27,28] while in mCL-t only backbone atoms are used.

Mechanism of Action

The mechanism of action for thermolysin as described by Matthews[27] has many features that are consistent with structural evidence found for mCL-t. A proposed mechanism for collagenase-catalyzed cleavage of peptides is shown in Figure 11. In mCL-t, Glu-219 appears to promote the nucleophilic

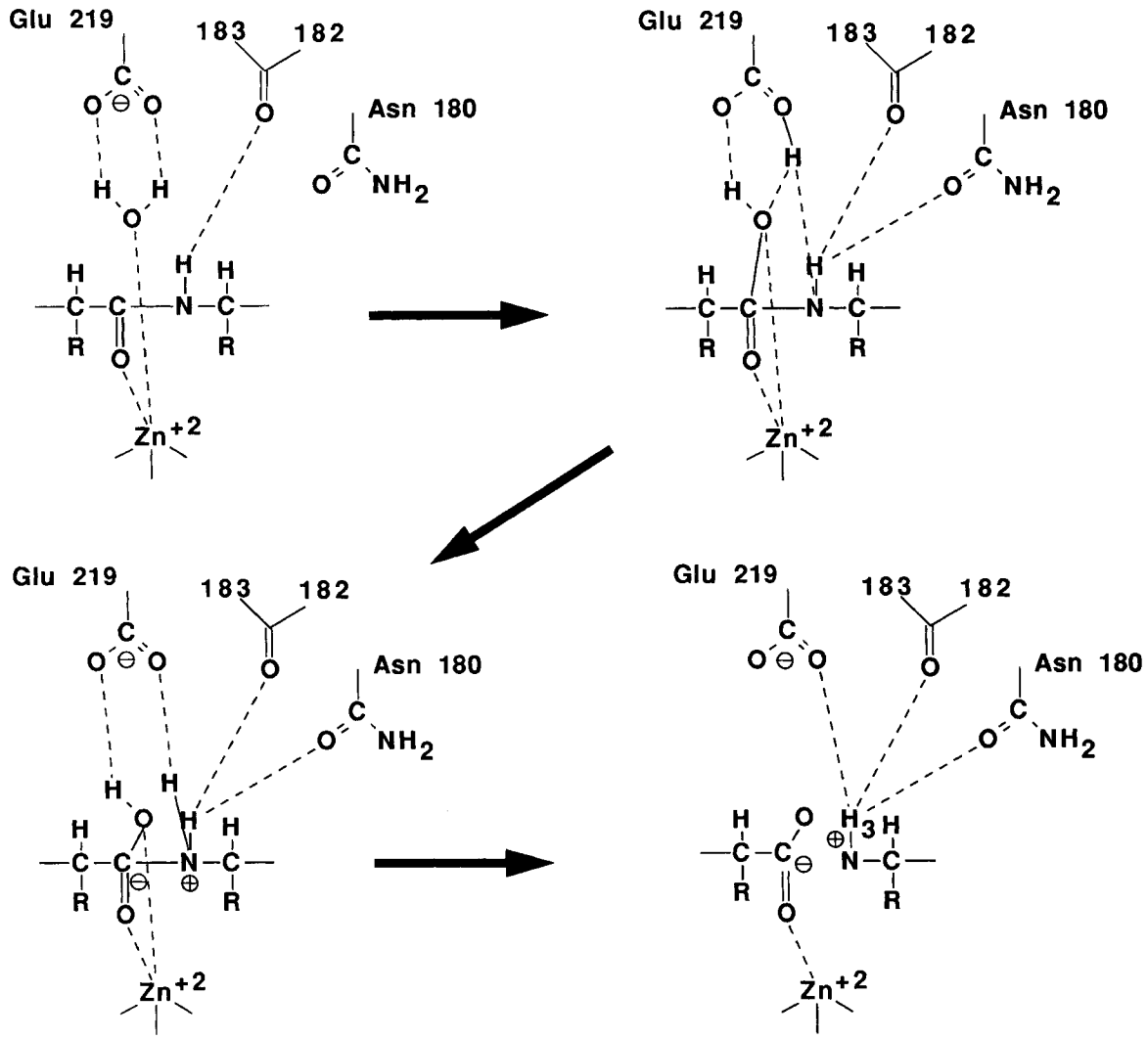

Fig. 11. A proposed mechanism for peptide cleavage by mCL-t.

attack of a water molecule on the scissile peptide bond, which is the same role Glu-143 plays in thermolysin. The carbonyl oxygen of residue 182 and the side chain oxygen of Asn-180 in mCL-t are proposed to stabilize the protonated nitrogen of the scissile bond, analogously to residues 113 and 112 in thermolysin.[27] No residues are seen in the truncated collagenase which could act to stabilize the carbonyl oxygen of the cleaved bond. The only stabilization appears to come from the carbonyl zinc interaction. In contrast thermolysin (and carboxypeptidase) utilizes a histidine and a tyrosine residue (or arginine) to stabilize the scissile carbonyl group.

The three-dimensional structure of mCL-t provides insight into the mechanisms for peptide cleavage by the collagenase family. The complete story on the method of collagen cleavage by these amazing enzymes, however, awaits the revelation of the structure of the full length mature enzyme.

ACKNOWLEDGMENTS

We thank Byron Rubin and Travis Stams for invaluable discussions and comments in the preparation of this manuscript.

REFERENCES

1. Woessner, F.J. Matrix metalloproteinases and their inhibitors in connective tissue remodeling. FASEB J. 5:2145–2154, 1991.
2. Muller, D., Quantin, B. Gesnel, M.-C., Millon-Collard, R., Abecassis, J., Breathnach, R. The collagenase gene family in humans consists of at least four members. Biochem. J. 253:187–192, 1988.
3. Emonard, H., Grimaud, J.-A. Matrix metalloproteinases. A review. Cell. Mol. Biol. 36:131–153, 1990.
4. Welgus, H.G., Jeffery, J.J., Eisen, A.Z. The collagen substrate specificity of human skin fibroblast collagenase. J. Biol. Chem. 256:9511–9515, 1981.
5. Vavra, K., Falvo, J., Banks, T., Schneider, E., Sochor, A., Roth, E., Lirette, R., Dixon, D., Brake, P., Pulvino, T., Ryerson, C., Ciccarelli, R., Wahl, R. Expression, purification and characterization of a C-terminally truncated recombi-

nant human fibroblast collagenase. Protein Sci., submitted.
6. Whitham, S.E., Murphy, G., Angel, P., Rahmsdorf, H.-J., Smith, B.J., Lyons, A., Harris, T.J.R., Reynolds, J.J., Herrlich, P., Docherty, A.J.P. Comparison of human stromelysin and collagenase by cloning and sequence analysis. Biochem. J. 240:913–916, 1986.
7. Lowry, C.L., McGeehan, G., Levine, H. Metal ion stabilization of the conformation of a recombinant 19-kDa catalytic fragment of human fibroblast collagenase. Proteins 12:42–48, 1992.
8. Howard, A.J., Nielson, C., Xoung, Ng. H. Software for a diffractometer with multi-wire area detector. Methods Enzymol. 114:452–472, 1985.
9. Howard, A.J., Gilliland, G.L., Finzel, B.C., Poulos, T.L., Ohlendorf, D.H., Salemme, F.R. The use of an imagining proportional counter in macromolecular crystallography. J. Appl. Crytallogr. 20:383–387, 1987.
10. Leslie, A.G.W. Recent changes to the MOSFLM package for processing film and image plate data. CCP4 and ESF-EACMB Newsletter on Protein Crystallography, Number 26, 1992.
11. Steigemann, W. Protein: A program system for the crystal structure analysis of proteins. Version 3.1. Users Guide (Max-planck-Institut für Biochemie, 1992).
12. Brünger, A.T. X-PLOR Version 3.1: A System for Crystallography and NMR. Reference manual. (Yale University, 1992).
13. Hendrickson, W.A. Stereochemically restrained refinement of macromolecular structures. Methods Enzymol. 115:252–270, 1985.
14. Sack, J.S. CHAIN—A crystallographic modeling program J. Mol. Graphics 6:244–245, 1988.
15. Sigler, P.B. Iodination of a single tyrosine in crystals of α-chymotrypsin. Biochemistry 9:3609–3617, 1970.
16. Stewart, D.E., Sarker, A., Wampler, J.E. Occurence and role of cis peptide bonds in protein structures. J. Mol. Biol. 214:253–260, 1990.
17. Carson, M. Ribbon models of macromolecules. J. Mol. Graphics 5:103–106, 1987.
18. Springman, E.B., Angleton, E.L., Birkedal-Hansen, H., Van Wart, H.E. Multiple modes of activation of latent human fibroblast collagenase: evidence for the role of a Cys^{73} active-site zinc complex in latency and a "cysteine switch" mechanism for activation. Proc. Natl. Acad. Sci. U.S.A. 87:364–368, 1990.
19. Holland, D.R., Tronrud, D.E., Pley, H.W., Flaherty, K.M., Stark, W., Jansonius, J.N., McKay, D.B., Mathews, B.W. Structural comparison suggests that thermolysin and related neutral proteases undergo hinge-bending motion during catalysis. Biochemistry 31:11310–11316, 1992.
20. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F., Brice, M.D., Rogers, J.R., Kennard, O., Schimanouchi, T., Tasumi, M. The protein data bank: A computer based archival file for macro-molecular structures. J. Mol. Biol. 122:535–542, 1977.
21. Rossmann, M.G., Argos, P.J. Exploring structural homology of proteins. Mol. Biol. 105:75–96, 1976.
22. Bode, W., Gomis-Rüth, F.X., Huber, R., Zwilling, R., Stöcker, W. Structure of astacin and implications for activation of astacins and zinc-ligation of collagenases. Nature (London) 358:164–167, 1992.
23. Gomis-Rüth, F.-X., Kress, L.F., Bode, W. First structure of a snake venom metalloproteinase: a prototype for matrix metalloproteinases/collagenases. EMBO J. 12:4151–4157, 1993.
24. Salowe, S.P., Marcy, A.I., Gregory, C.C., Smith, C.K., Kopka, I.E., Hagmann, W.K., Hermes, J.D. Characterization of zinc-binding sites in human stromelysin-1: Stoichiometry of the catalytic domain and the identification of a cysteine ligand in the proenzyme. Biochemistry 31:4535–4540, 1992.
25. Holmes, M.A., Mathews, B.W. Binding of hydroxamic acid inhibitors to crystalline thermolysin suggests a pentacoordinate zinc intermediate in catalysis. Biochemistry 20: 6912–6920, 1981.
26. Ponder, J.W., Richards, F.M. Tertiary templates for proteins. J. Mol Biol. 193:775–791, 1987.
27. Matthews, B.W. Structural basis of the action of thermolysin and related zinc peptidases. Acc. Chem. Res. 21:333–340, 1988.
28. Monzingo, A.F., Matthews, B.W. Binding of N-carboxymethyl dipeptide inhibitors to thermolysin determined by x-ray crystallography: A novel class of transition-state analogues for zinc peptidases. Biochemistry 23:5724–5729, 1984.