Structure of the Catalytic Domain of Fibroblast Collagenase Complexed with an Inhibitor

Brett Lovejoy, Anne Cleasby, Anne M. Hassell, Kelly Longley, Michael A. Luther, Debra Weigl, Gerard McGeehan, Andrew B. McElroy, David Drewry, Millard H. Lambert, Steven R. Jordan

Collagenase is a zinc-dependent endoproteinase and is a member of the matrix metalloproteinase (MMP) family of enzymes. The MMPs participate in connective tissue remodeling events and aberrant regulation has been associated with several pathologies. The 2.4 angstrom resolution structure of the inhibited enzyme revealed that, in addition to the catalytic zinc, there is a second zinc ion and a calcium ion which play a major role in stabilizing the tertiary structure of collagenase. Despite scant sequence homology, collagenase shares structural homology with two other endoproteinases, bacterial thermolysin and crayfish astacin. The detailed description of protein-inhibitor interactions present in the structure will aid in the design of compounds that selectively inhibit individual members of the MMP family. Such inhibitors will be useful in examining the function of MMPs in pathological processes.

The MMPs are responsible for cleaving extracellular matrix components during normal connective tissue remodeling events such as embryonic development and wound healing. Members of the MMP family include collagenases, stromelysins, and gelatinases (*1*, *2*). These enzymes have broad, overlapping substrate specificities and are regulated by a wide range of factors such as growth factors, hormones, and the tissue inhibitor of metalloproteinases (TIMP) family of proteins (*1*, *2*). MMP activity has also been associated with pathologies such as arthritis and metastatic cancer (*3*, *4*). However, the exact role of specific MMPs in physiological and pathological processes has not been established. To help determine their function, there have been efforts to develop compounds that selectively inhibit individual members of the MMP family (*5*).

The MMPs have an NH_2-terminal propeptide (~80 residues) that contains a free cysteine thought to maintain enzyme latency by ligating to the catalytic zinc. The propeptide is cleaved to form the mature, active enzyme (*1*, *2*). The active enzyme contains a zinc- and calcium-binding catalytic domain (~180 residues) and a COOH-terminal domain (~200 residues) that may be involved in matrix binding (*6*) and alignment of the catalytic domain at the site of enzymatic cleavage (*7*). The catalytic domains of MMPs share the highest sequence homology. Recombinant collagenase constructs containing only the catalytic domain can cleave casein, gelatin, and peptide substrates but, unlike the full-length enzyme, cannot cleave collagen (*7*, *8*). With peptide substrates, however, collagenase inhibitors exhibit similar inhibitory activity toward the whole enzyme as they do on the catalytic domain alone (*8*). Previous models of the MMP catalytic domain were based on the structures of the bacterial endoproteinase, thermolysin, complexed with inhibitors (*9*). Recently the structure of astacin, an endoproteinase that is more closely related to MMPs, has been reported (*10*). Thermolysin- and astacin-based models of MMPs have not provided the detail necessary to design highly selective MMP inhibitors. Here we report the x-ray crystal structure of the catalytic domain of human fibroblast collagenase (residues 101 to 269) (*11*, *12*) bound to a carboxyalkylamine-based inhibitor (Fig. 1A).

The catalytic domain of human fibroblast collagenase was solved at 2.4 Å resolution by multiple isomorphous replacement (MIR) (Fig. 1B) and refined to a crystallographic *R* factor of 0.186 (Table 1). The structure (Fig. 2A) consists of a twisted five-stranded β sheet and three long α helices (A, B, and C) (*13*). Collagenase shares structural homology with the NH_2-terminal domain of astacin (*10*) and, to a lesser extent, the NH_2-terminal domain of thermolysin (*14*). As with astacin, the β sheet contains four parallel strands (I, II, III, and V) and one antiparallel strand (IV). The collagenase active site cleft is bordered by β strand IV, helix B, and a stretch of random coil adjacent to the COOH-terminus of helix B. The catalytic zinc is at the bottom of the cleft and is ligated by His^{218}, His^{222}, and His^{228}, each at a distance of 2.1 Å (Fig. 2A). The collagenase catalytic zinc environment is similar to that in astacin, with the first two zinc ligands, His^{218} and His^{222} (His^{92} and His^{96} in astacin), located on a central α helix (helix B in both collagenase and astacin) that extends up to Gly^{225} (Gly^{99} in astacin). Here the chain turns sharply toward the third zinc ligand, His^{228} (His^{102} in astacin). The carboxylate group of the inhibitor provides an additional ligand to the collagenase catalytic zinc with Oε1 and Oε2

Table 1. Data collection and phasing statistics. Native crystals consist of the catalytic domain of collagenase cocrystallized with an inhibitor. They are in space group $P6_4$ with cell dimensions $a = b = 78.2$ Å and $c = 87.4$ Å, and there are two molecules in the asymmetric unit (*21*). Collagenase was determined with MIR data collected on two derivatives (*22*). Each data set was obtained from one crystal with a Rigaku R-Axis imaging plate. Heavy atom parameters were refined and phases calculated with PHASES (*23*); the mean figure of merit was 0.50 (30 to 2.5 Å). Initial MIR phases were solvent flattened, resulting in an improved electron density map (mean figure of merit 0.87, 30 to 2.5 Å) into which molecule A (residues 106 to 268) and molecule B (residues 106 to 152 and 158 to 265) were traced. The electron density maps contained three peaks associated with each polypeptide: Two were identified as zinc ions (10σ and 7σ), and one was identified as a calcium ion (6σ). Several cycles of simulated annealing (*24*), manual model rebuilding, and atomic temperature factor refinement resulted in a final model that includes 65 water molecules, molecule A (residues 101 to 268), and molecule B (residues 105 to 152 and 158 to 265). The final model has an *R* factor of 18.6% for all $I/\sigma(I) \geq 1.0$ data (10,399 unique reflections) from 7 to 2.4 Å resolution. Root-mean-square (rms) deviations from ideal bonds and angles were 0.012 Å and 2.2°, respectively. The rms deviation between the Cα carbons of molecules A and B (157 positions) is 0.38 Å. No residue φ,ψ angles lie outside allowed regions of the Ramachandran plot.

Data set	Resolution (Å)	Unique reflections	R_{merge} (%)	R_{scale}† (%)	Number of sites	Phasing power‡	R_{Cullis} (30 to 2.5 Å)§
Native	2.4	10,896	6.7				
Iodine	2.4	8,778	7.6	17.7	2	1.84	0.50
Iodine·$Pb(CH_3COO)_2$	2.5	8,782	8.2	24.8	5‖	1.88	0.44

*$R_{merge} = \Sigma_i|\langle I\rangle - I_i|/\Sigma_i I_i$ and $\langle I\rangle$ is the average of I_i over all symmetry equivalents. †$R_{scale} = \Sigma|F_{PH} - F_P|/\Sigma F_P$, where F_{PH} and F_P are the derivative and native structure factors, respectively. ‡Phasing power is the mean value of the heavy atom structure factor amplitude divided by the residual lack of closure error. §$R_{Cullis} = \Sigma|F_{PH} - F_P - f_H|/\Sigma|F_{PH} - F_P|$ for centric reflections, where f_H is a calculated heavy atom structure factor. ‖There are two iodine sites and three lead sites.

B. Lovejoy, A. M. Hassell, K. Longley, M. A. Luther, D. Weigl, G. McGeehan, A. B. McElroy, D. Drewry, M. H. Lambert, S. R. Jordan, Glaxo Research Institute, Research Triangle Park, NC 27709.
A. Cleasby, Glaxo Group Research, Greenford, Middlesex UB 06 HE, United Kingdom.

at distances of 2.0 and 2.6 Å, respectively.

The inhibitor is tightly bound within the collagenase active site cleft with an inhibition constant of 135 nM (*15*). The complex appears to be stabilized by the ligation of the inhibitor carboxylate group to the catalytic zinc, as well as through hydrogen bonds and hydrophobic interactions. Catalytic zinc ligation by an inhibitor carboxylate group has also been observed in thermolysin-inhibitor complexes. Superposition of one such complex, described by Monzingo and Matthews (*16*), onto the collagenase-inhibitor complex reveals that the carboxylate groups ligate the catalytic zinc in a similar manner (Fig. 3A). The collagenase inhibitor forms eight hydrogen bonds with the enzyme (*17*) (Fig. 3B). Several of these bonds are analogous to bonds found in the thermolysin-inhibitor complex of Monzingo and Matthews (*16*). Both Glu^{219} Oε1 and the corresponding thermolysin residue, Glu^{143}, interact with the P1′ backbone nitrogen atom of their respective inhibitors. In both complexes, the P1′ nitrogen mimics the nitrogen of the scissile bond, and in thermolysin, Glu^{143} is postulated to donate a proton to that nitrogen (*9*). In some thermolysin-inhibitor complexes, Asn^{112} also forms a hydrogen bond with the P1′ nitrogen (*9, 16*). In collagenase, the corresponding Asn^{180} makes a hydrogen bond with the P1 nitrogen and carbonyl oxygen instead. Hydrophobic interactions found in the collagenase-inhibitor complex are comparable with those found in thermolysin-inhibitor complexes as well. For example, the P1′ leucine side chain of the collagenase inhibitor and the corresponding P1′ leucine side chain of several thermolysin inhibitors (*9, 16*) are directed into analogous hydrophobic pockets (Fig. 3A).

Structural studies of thermolysin-inhibitor complexes suggest that Tyr^{157} and His^{231} stabilize the negative charge that forms in the transition state of the thermolysin reaction mechanism (*9, 16*). There are no residues comparable with Tyr^{157} or His^{231} in the collagenase catalytic domain. It is possible that hydrogen-bond donors come from water molecules or the COOH-terminal domain. However, naturally occurring MMPs that have no COOH-terminal domain exist (*18, 19*). The collagenase catalytic zinc has a higher net positive charge than the thermolysin catalytic zinc because it is ligated by three neutral histidines, whereas one of the ligands in thermolysin is a negatively charged carboxylate (Fig. 3A). As a result, the collagenase catalytic zinc may play a larger role in stabilizing the negative charge in the transition state and reduce the need for transition state hydrogen bond donors. The histidines that ligate the collagenase catalytic zinc are conserved in all MMPs (*1, 2, 18*).

In addition to the catalytic zinc, a second zinc ion has been identified by its relative intensity in electron density maps (Table 1) and the nature of its ligands. The second zinc ion is coordinated by His^{168} Nε2 (2.0 Å), Asp^{170} Oδ2 (2.2 Å), His^{183} Nε2 (2.2 Å), and His^{196} Nδ1 (2.0 Å) in an approximately tetrahedral manner. The presence of a second zinc ion in collagenase is consistent with biochemical studies of stromelysin-1 that have identified two zincs bound to the catalytic

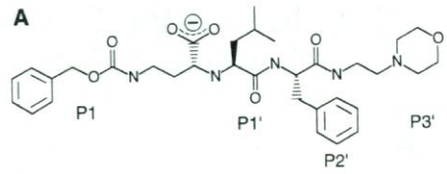
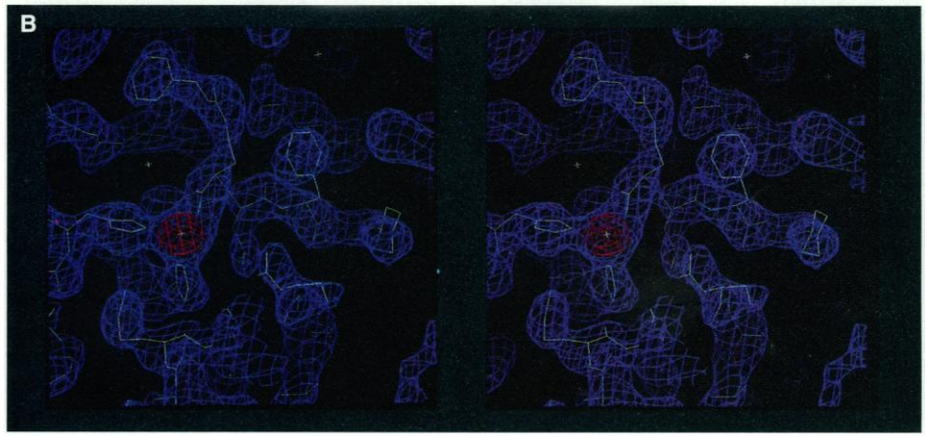


Fig. 1. (**A**) Structure of the collagenase inhibitor. The central carboxylate group binds the active site zinc. P1′ is leucine, P2′ is phenylalanine, P3′ is a morpholino group, and P1 is a benzyloxycarbonylamino group. (**B**) A stereo view of the MIR electron density map (contoured at 1σ) at 2.5 Å resolution showing the enzyme-inhibitor complex. The active site zinc is displayed as a red surface.



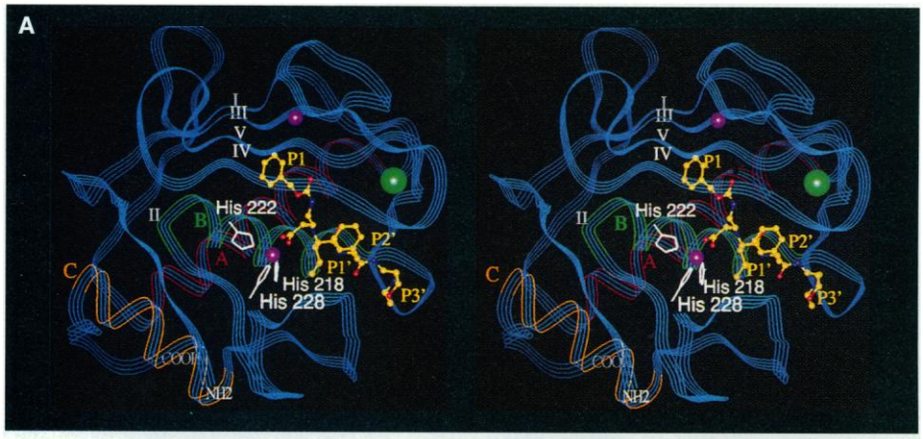
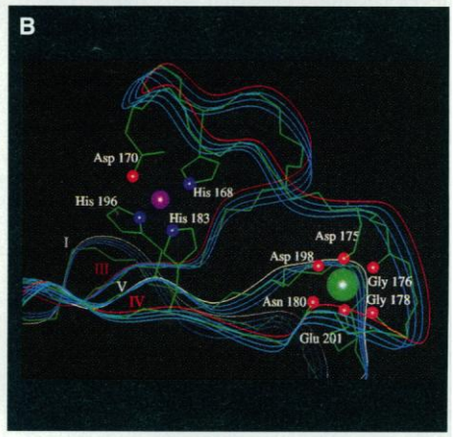


Fig. 2. (**A**) A stereo Cα ribbon model of the collagenase-inhibitor complex. The inhibitor is displayed with yellow carbon atoms, red oxygen atoms, and navy blue nitrogen atoms. The catalytic zinc (lower left), secondary zinc (upper left), and calcium (right) ions are also shown. The β strands are labeled in white, helix A is red, helix B is green, and helix C is orange. The residues His^{218}, His^{222}, and His^{228} ligate the catalytic zinc. (**B**) A ribbon diagram shows how the second zinc ion (left) and the calcium ion (right) fasten the extended loop between β strands III and IV onto the β sheet. Atoms that ligate the zinc or calcium ions are shown as spheres. The β strands III and IV and the extended loop contain a red strand. Zinc ions are displayed as pink spheres and the calcium ion is displayed as a green sphere in both panels.



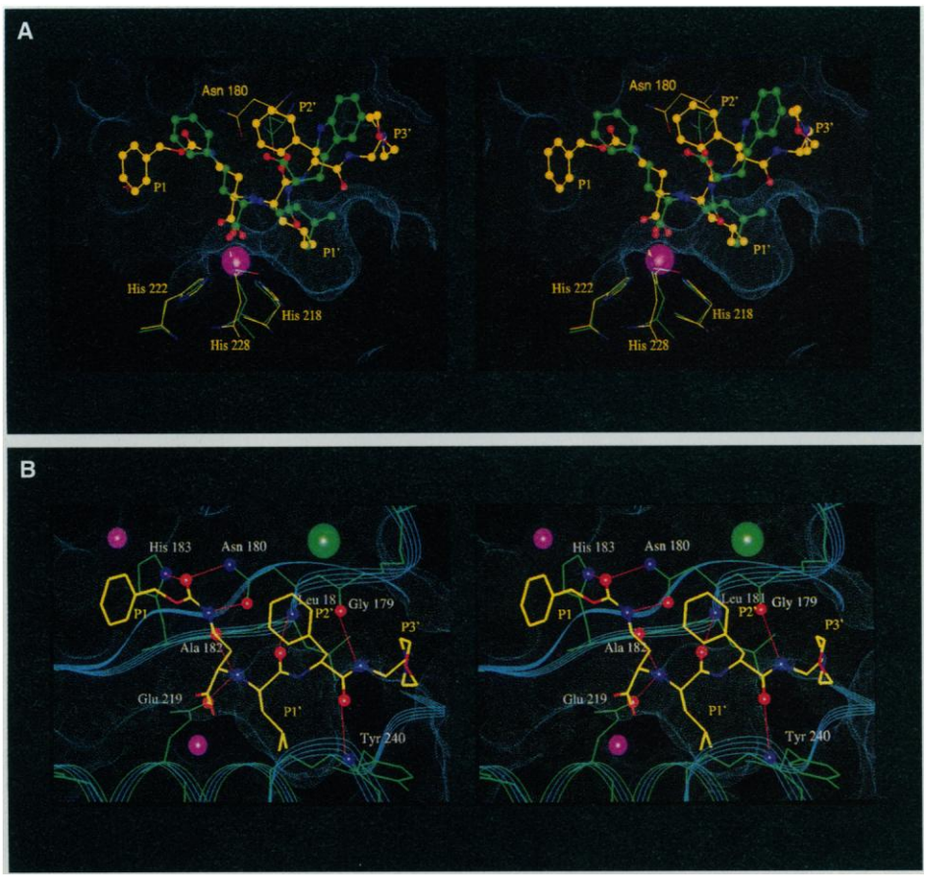


Fig. 3. (**A**) A stereo view of the binding mode of the collagenase inhibitor compared with that of a thermolysin inhibitor (*16, 25*). Oxygen and nitrogen atoms are red and navy blue, respectively. Carbon atoms are yellow in the collagenase-inhibitor complex and green in the thermolysin-inhibitor complex. Atoms in the inhibitors are represented by solid spheres. (**B**) A stereo view of the collagenase inhibitor (yellow) bound within the active site cleft. A light blue ribbon traces the Cα positions of neighboring collagenase residues including helix B (green). The catalytic zinc (lower left), second zinc (upper left), and calcium (right) ions are also shown. Collagenase residues are green except for atoms that form hydrogen bonds with the inhibitor. Such atoms are drawn as navy blue (nitrogen) or red (oxygen) spheres. The eight enzyme-inhibitor hydrogen bonds (*17*) are represented by red lines. In addition to forming a hydrogen bond with the inhibitor, His^{183} is ligated to the second zinc ion (upper left). The collagenase active site cleft is mapped with a blue Connolly surface (with a rolling sphere of 1.4 Å radius) in both panels.

domain (*20*). Sequence alignment (*1, 18*) of the catalytic domains of the MMPs shows that these four ligands are completely conserved within the enzyme family, suggesting that the presence of two zinc ions is a common structural feature of MMP catalytic domains.

The calcium ion present in the complex is ligated in an octahedral manner by Asp^{175} Oδ1 (2.2 Å), Asp^{198} Oδ2 (1.9 Å), Glu^{201} Oε2 (1.8 Å), and the main chain carbonyl oxygens of Gly^{176} (2.2 Å), Gly^{178} (2.5 Å), and Asn^{180} (2.2 Å). The contribution of four of the calcium ligands from the sequence Asp-Gly-Pro-Gly-Gly-Asn (residues 175 to 181) tightly orders this sequence. By binding both the β sheet and the large, highly exposed loop (residues 165 to 181) between strands III and IV, the second zinc ion and the calcium ion effectively fasten the loop onto the β sheet (Fig. 2B). Earlier studies demonstrated that zinc and calcium significantly stabilize the tertiary structure of collagenase (*12*). These data suggest that the interaction of the exposed loop with the second zinc ion and the calcium ion is essential for maintaining the structural integrity of collagenase.

The role of MMPs in a wide array of physiological and pathological processes is complex and largely undefined. The structure of collagenase complexed to an inhibitor will facilitate the detailed modeling of other MMPs. These models will aid in the structure-based design of compounds that selectively inhibit individual members of the MMP family. Such inhibitors will be useful for examining the function of MMPs and may have therapeutic value.

REFERENCES AND NOTES

1. H. Birkedal-Hansen *et al.*, *Crit. Rev. Oral Biol. Med.* **4**, 197 (1993).
2. J. F. Woessner Jr., *FASEB J.* **5**, 2145 (1991).
3. S. S. McCachren, *Arthritis Rheum.* **34**, 1085 (1991).
4. L. A. Liotta, C. N. Rao, U. M. Wewer, *Annu. Rev. Biochem.* **55**, 1037 (1986).
5. B. Henderson, A. J. P. Docherty, N. R. A. Beeley, *Drugs Future* **15**, 495 (1990).
6. J. A. Allan, R. M. Hembry, S. Angal, J. J. Reynolds, G. Murphy, *J. Cell Sci.* **99**, 789 (1991).
7. G. Murphy *et al.*, *J. Biol. Chem.* **267**, 9612 (1992).
8. J. D. Becherer *et al.*, *J. Cell. Biochem. Suppl.* **15G** (1991).
9. B. W. Mathews, *Acc. Chem. Res.* **21**, 333 (1988).
10. W. Bode, F. X. Gomis-Rüth, R. Huber, R. Zwilling, W. Stöcker, *Nature* **358**, 164 (1992); F. X. Gomis-Rüth, W. Stöcker, R. Huber, R. Zwilling, W. Bode, *J. Mol. Biol.* **229**, 945 (1993).
11. G. I. Goldberg *et al.*, *J. Biol. Chem.* **261**, 6600 (1986).
12. C. L. Lowry, G. McGeehan, H. LeVine III, *Proteins Struct. Funct. Genet.* **12**, 42 (1992).
13. The secondary structure consists of helices A (residues 126 to 143), B (211 to 225), and C (249 to 260) and β strands I (residues 112 to 118), II (146 to 151), III (158 to 164), IV (181 to 186), and V (193 to 198).
14. B. W. Matthews, J. N. Jansonius, P. M. Colman, B. P. Schoenborn, D. Dupourque, *Nat. New Biol.* **238**, 37 (1972).
15. F. K. Brown *et al.*, *J. Med. Chem.*, in press; G. M. McGeehan, unpublished results.
16. A. F. Monzingo and B. W. Matthews, *Biochemistry* **23**, 5724 (1984).
17. In molecule A, P1 N and CO form hydrogen bonds with Asn^{180} Oδ1 and Nδ2 (lengths 3.0 and 2.7 Å, respectively), P1 CO hydrogen bonds to His^{183} Nδ1 (2.7 Å), P1′ N hydrogen bonds to Ala^{182} CO and Glu^{219} Oε1 (3.0 and 3.2 Å, respectively), P1′ CO hydrogen bonds to Leu^{181} N (2.8 Å), P2′ CO hydrogen bonds to Tyr^{240} N (3.0 Å), and the backbone N of P3′ hydrogen bonds to Gly^{179} CO (3.1 Å). Distances in molecule B do not differ significantly.
18. G. McGeehan *et al.*, *Plant Physiol.* **99**, 1179 (1992).
19. D. Muller *et al.*, *Biochem. J.* **253**, 187 (1988).
20. S. P. Salowe *et al.*, *Biochemistry* **31**, 4535 (1992).
21. Crystals were obtained from 15 to 30% (w/v) polyethylene glycol 8000 and 0.2 M ammonium sulfate in 0.1 M cacodylate buffer, pH 6.5 (A. M. Hassell *et al.*, *J. Mol. Biol.*, in press).
22. The iodine derivative was obtained by cocrystallizing collagenase with an inhibitor analog that contained an iodine at the para position of the P2′ ring. The iodine-lead derivative was obtained by soaking the crystal used to collect the iodine derivative in 5 mM $Pb(CH_3COO)_2$ in an artificial mother liquor [40% (w/v) polyethylene glycol 8000 in 0.1 M cacodylate buffer, pH 6.5] for 4 days. Iodine sites were determined by inspection of difference Patterson maps, whereas lead sites were determined by inspection of difference Fourier maps with the iodine sites for phasing.
23. W. Furey and S. Swaminathan, *Am. Crystallogr. Assoc. Annu. Mtg. Program Abst.* **18**, 73 (1990).
24. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987). The Lennard-Jones σ parameter for zinc, but not calcium, was adjusted to obtain distances similar to those found in small molecule crystal structures.
25. Thermolysin was superimposed onto collagenase by a least squares fit of His^{142}, His^{146}, and the thermolysin catalytic zinc onto His^{218}, His^{222}, and the collagenase catalytic zinc, respectively, with the method of W. Kabsch [*Acta Crystallogr. Sect. A* **34**, 827 (1978)].

26. We thank J. Veal and F. Brown for helpful discussions and P. Graber, E. Sebille, A. Bernard, and T. Wells for assistance in expressing collagenase. Coordinates have been deposited with the Brookhaven Protein Database. Figures 2 and 3 were prepared with Insight II (Biosym Technologies, Incorporated, San Diego, CA).

9 September 1993; accepted 9 December 1993