Protein Science (1995), *4*:1966–1976. Cambridge University Press. Printed in the USA.
Copyright © 1995 The Protein Society

Stromelysin-1: Three-dimensional structure of the inhibited catalytic domain and of the C-truncated proenzyme

JOSEPH W. BECKER,[1] ALICE I. MARCY,[1] LAURA L. ROKOSZ,[1,3] MELINDA G. AXEL,[1] JONATHAN J. BURBAUM,[1,3] PAULA M.D. FITZGERALD,[1] PATRICIA M. CAMERON,[1] CRAIG K. ESSER,[2] WILLIAM K. HAGMANN,[2] JEFFREY D. HERMES,[1] AND JAMES P. SPRINGER[1]

[1] Department of Molecular Design and Diversity, and [2] Department of Medicinal Chemical Research, Merck Research Laboratories, Rahway, New Jersey 07065

(RECEIVED May 18, 1995; ACCEPTED August 1, 1995)

Abstract

The proteolytic enzyme stromelysin-1 is a member of the family of matrix metalloproteinases and is believed to play a role in pathological conditions such as arthritis and tumor invasion. Stromelysin-1 is synthesized as a proenzyme that is activated by removal of an N-terminal prodomain. The active enzyme contains a catalytic domain and a C-terminal hemopexin domain believed to participate in macromolecular substrate recognition. We have determined the three-dimensional structures of both a C-truncated form of the proenzyme and an inhibited complex of the catalytic domain by X-ray diffraction analysis. The catalytic core is very similar in the two forms and is similar to the homologous domain in fibroblast and neutrophil collagenases, as well as to the stromelysin structure determined by NMR. The prodomain is a separate folding unit containing three α-helices and an extended peptide that lies in the active site of the enzyme. Surprisingly, the amino-to-carboxyl direction of this peptide chain is opposite to that adopted by the inhibitor and by previously reported inhibitors of collagenase. Comparison of the active site of stromelysin with that of thermolysin reveals that most of the residues proposed to play significant roles in the enzymatic mechanism of thermolysin have equivalents in stromelysin, but that three residues implicated in the catalytic mechanism of thermolysin are not represented in stromelysin.

Keywords: drug design; matrix metalloproteinase; X-ray-crystallography; zymogen activation

Stromelysin-1 (EC 3.4.27.17) is a proteolytic enzyme and a member of the matrix metalloproteinase (MMP) family. These proteins represent attractive targets for drug development because their ability to degrade the protein components of connective tissue is believed to underlie fundamental events in both rheumatoid arthritis and osteoarthritis. In addition, there is evidence that this activity also plays a role in tumor invasion (Murphy et al., 1991; Woessner, 1991; Docherty et al., 1992; Birkedal-Hansen et al., 1993). The MMPs comprise three broad families, defined by their substrate specificity: the collagenases, which digest interstitial collagen (Welgus et al., 1981; Schmid et al., 1986; Seltzer et al., 1989), the gelatinases, which digest denatured collagens and gelatins (Murphy et al., 1985; Seltzer et al., 1989), and the stromelysins, which display a broader specificity (Galloway et al., 1983; Chin et al., 1985; Okada et al., 1989). This broad specificity, together with the observations that they are induced by the inflammatory mediator interleukin-1 (Gowen et al., 1984; Saus et al., 1988; MacNaul et al., 1990) and can activate other members of the MMP family (Murphy et al., 1987; He et al., 1989; Ogata et al., 1992; Knäuper et al., 1993), suggests that the stromelysins may play a particularly significant role in the physiological processes mediated by this family of enzymes.

The MMPs show structural similarity to other metalloproteinases such as thermolysin and the astacin family of zinc endoproteinases. MMPs are synthesized as inactive precursors and comprise three distinct structural domains: an N-terminal propeptide, a catalytic domain of approximately 180 amino acid residues, and a C-terminal domain that appears to play a role in recognition of macromolecular substrates and interaction with macromolecular inhibitors. Active enzyme is produced by cleavage of the propeptide that can be induced by heat, mercurial re-

Reprint requests to: Joseph W. Becker, Merck Research Laboratories, RY50-105, P.O. Box 2000, Rahway, New Jersey 07065-0900; e-mail: beckerj@merck.com.

[3] Present address: Pharmacopeia, Inc., 101 College Road East, Princeton, New Jersey 08540.

agents, proteolytic enzymes, or nitric oxide (Okada et al., 1988; Okada & Nakanishi, 1989; Nagase et al., 1990; Koklitis et al., 1991; Murrell et al., 1995). The enzymatic activity, specificity, and sensitivity to inhibitors of the catalytic domain of stromelysin are similar to that of the full-length protein (Marcy et al., 1991; Ye et al., 1994). This finding has established that truncated stromelysin is an appropriate model for the structure-based design of inhibitors of the full-length protein. In addition, use of the truncated enzyme has facilitated structural studies on stromelysin because it precludes the structural heterogeneity caused by autolysis in the C-terminal domain of the full-length protein (Okada et al., 1986, 1988, 1989; Nagase et al., 1990).

Several three-dimensional structures of MMPs have recently been published. The structure of the inhibited catalytic domain of human stromelysin-1 was determined recently by multidimensional NMR methods (Gooley et al., 1994), and the structures of inhibited human fibroblast (Borkakoti et al., 1994; Lovejoy et al., 1994a, 1994b; Spurlino et al., 1994) and neutrophil (Bode et al., 1994; Stams et al., 1994) collagenases have been determined by X-ray diffraction analysis. These structures show considerable similarity to one another and belong to the "metzincin" family of proteins, which is characterized by an HExxHxxGxxH sequence motif, and includes digestive enzymes, snake venom metalloproteases, and bacterial proteases as well as MMPs (Bode et al., 1993; Gomis-Rüth et al., 1994; Stöcker et al., 1995). This family of proteins is also structurally related to bacterial metalloenzymes such as thermolysin (Matthews, 1988) that contain a related sequence motif, HExxH(~20x)NExSD.

Here, we report the three-dimensional structure of two forms of stromelysin-1: the C-truncated proenzyme and the complex of the catalytic domain with the N-carboxyalkyl peptide inhibitor **I** (Fig. 1) used in the previously reported NMR structure (Gooley et al., 1994). The structure of the catalytic domain of the proenzyme is remarkably similar to that of the active form, and the pro-domain is a separate folding unit characterized by three α-helices and an extended propeptide that occupies the active site and blocks the catalytic zinc ion. Surprisingly, the N-to-C direction of the propeptide chain is the reverse of that adopted by all peptide-based inhibitors observed in complexes of stromelysin and collagenase to date. The active site contains a groove in the protein surface and an extraordinarily large S1′ specificity pocket that extends through the full width of the catalytic domain. The active site of stromelysin contains a basic structure similar to that of thermolysin but lacks residues equivalent to three groups believed to play key roles in the mechanism of that enzyme.

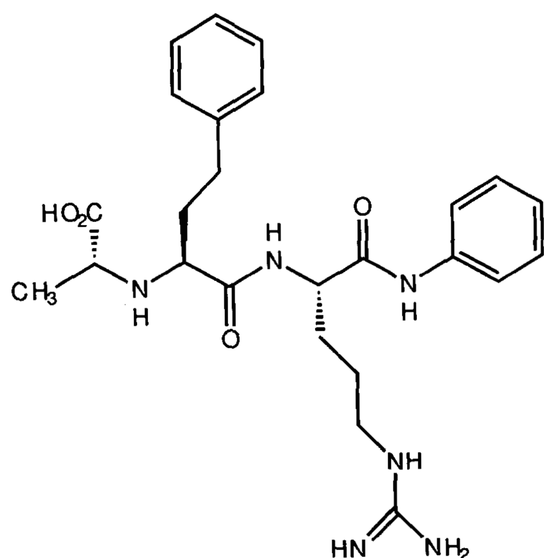

Fig. 1. N-carboxyalkyl peptide inhibitor (**I**, $K_i = 0.23$ μM [Chapman et al., 1993]) used to inhibit the catalytic domain of stromelysin.

Results

The structure of the inhibited catalytic domain of stromelysin determined in this study is similar to the solution structure determined by multidimensional NMR (Gooley et al., 1994). The protein is folded into a single globular unit approximately 35 Å in diameter, and the folding is dominated by a single five-stranded β-sheet, with one antiparallel and four parallel strands, and three α-helices (Fig. 2A). The propeptide (residues 16–82) makes up a separate smaller domain, approximately 20 Å in diameter and containing three α-helices (Fig. 2B). No electron density is visible for residues 1–15 and 31–39 of the prodomain. The catalytic domain contains two tetrahedrally coordinated Zn^{2+} ions: a "structural" zinc ion whose ligands are the side chains of Asp 153, His 151, His 166, and His 179, and a "catalytic" zinc ion whose ligands include the side chains of His 201, His 205, and His 211. In the inhibited complex, the fourth ligand of the catalytic zinc is the carboxylate group of the inhibitor; in the proenzyme, it is the sulfur atom of Cys 75. Two electron densities with roughly octahedral coordination, apparently Ca^{2+} ions, are also present in both the complex and the proenzyme. The first calcium site is defined by the ligands Asp 158 $O^{\delta 1}$, Gly 159 O, Gly 161 O, Val 163 O, Asp 181 $O^{\delta 2}$, and Gly 184 $O^{\epsilon 2}$. The second site includes Asp 141 O, Gly 173 O, Asn 175 O, Asp 177 $O^{\delta 1}$, and two water molecules, one of which is also hydrogen bonded to Gly 171 O. In the inhibited complex, there is an apparent third Ca^{2+} site defined by Asp 107 $O^{\delta 2}$, Asp 182 O, Asp 182 $O^{\delta 1}$, and Glu 184 O. This site is probably only partially occupied: the temperature factor is 19.4 $Å^2$, significantly higher than those for the other metal ions (3.1–10.1 $Å^2$), and there is only weak electron density at the two remaining octahedral ligand sites, consistent with partial occupancy by water. In the proenzyme structure, this site apparently is occupied by a water molecule, or possibly by a very low occupancy calcium ion: attempts to refine a calcium ion produced very high temperature factors, and refinement with a water molecule results in a temperature factor of 21.2 $Å^2$, similar to the average temperature factor for the other ligand atoms, 24.6 $Å^2$.

The active site consists of two distinct regions: a groove in the protein surface, centered on the catalytic zinc (Fig. 3), and a large, predominantly hydrophobic, S1′ site that extends completely through the body of the molecule (Figs. 4, 5). In both the proenzyme and the inhibited protein, the groove is occupied by extended peptide chains that make several β-structure-like hydrogen bonds with the enzyme and provide the fourth ligand for the catalytic zinc ion (Fig. 3). Despite the fact that the propeptide and the inhibitor make similar interactions with the same groups of the enzyme, the direction of the polypeptide chain is different in the two structures (Fig. 3). The S1′ subsite is empty in the proenzyme structure but contains the side chain of the

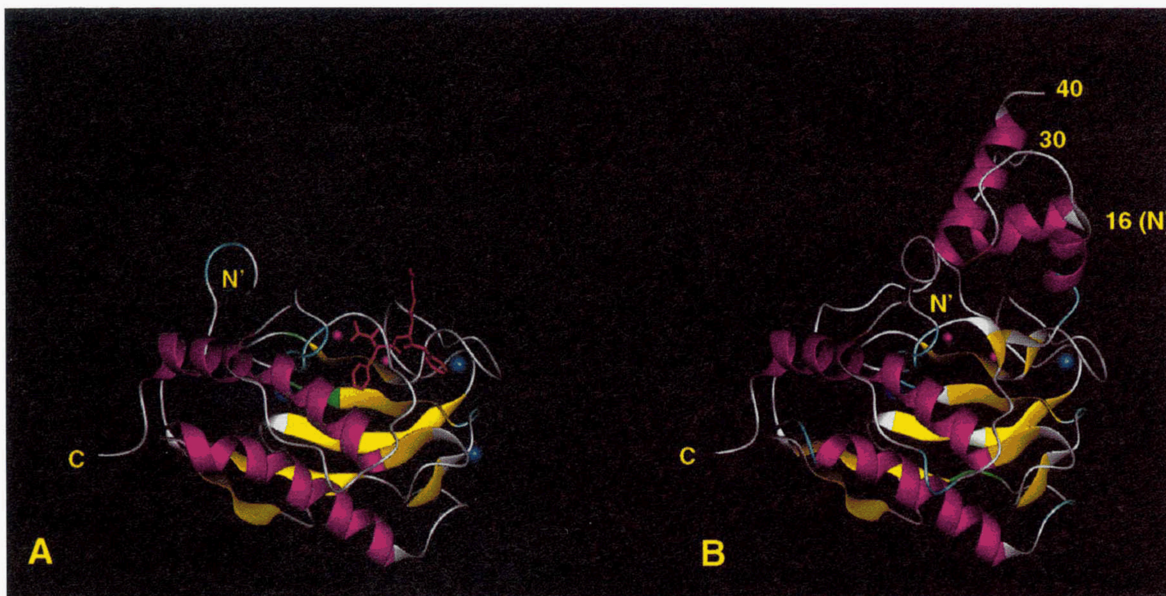

Fig. 2. A: Ribbon drawing of the stromelysin complex with **I**. Zinc and calcium ions are represented by violet and light blue spheres, respectively. Bound inhibitor is shown in red. The N-terminus of the mature protein is indicated by N′, the C-terminus by C. **B:** Ribbon drawing of prostromelysin. The first visible residue in the polypeptide chain, Leu 16, is labeled. Residue numbers of other notable features in the pro-domain are indicated.

homophenylalanyl residue in the inhibited complex. The aromatic ring is surrounded by hydrophobic moieties punctuated by two carbonyl groups pointed directly at its edge (Fig. 4). The side chain of His 201, a ligand of the catalytic zinc ion, is parallel to the ring. From this point, residues 218–223 make up a cordon that encircles half of the homophenylalanyl ring. The carbonyl oxygen atoms of Leu 218 and Tyr 220 point directly at adjacent carbon atoms on one side of the ring, and the edge of the side chain of Tyr 223 is directed toward the center of the ring, approximately 180° from the side chain of His 201. The

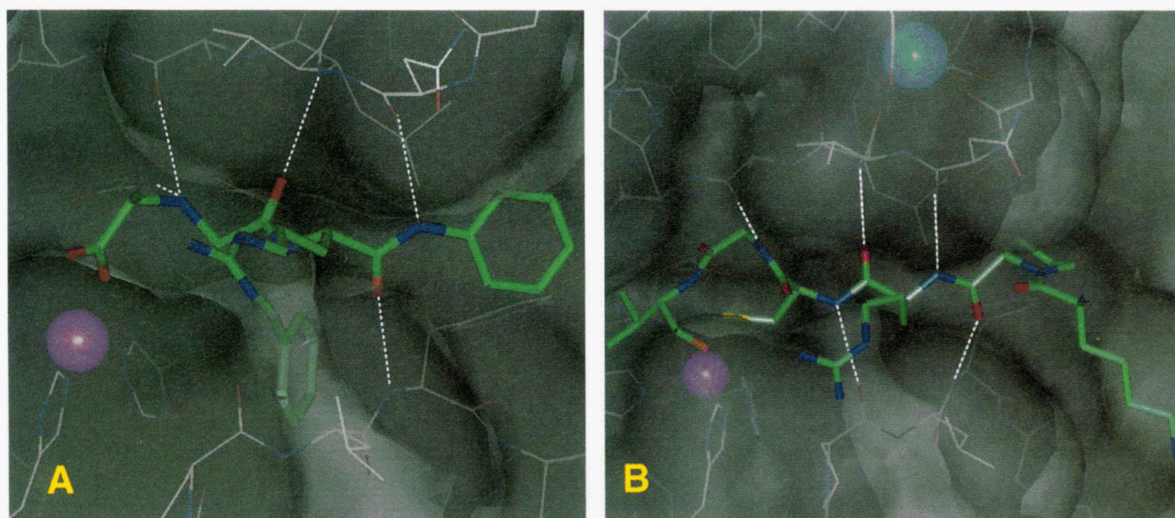

Fig. 3. A: Active site groove of the complex with **I**. Hydrogen bonds between the bound inhibitor and the protein are indicated by dashed lines. Violet sphere represents the catalytic zinc ion. The large S1′ pocket is at the bottom center of the figure. The orientation of the figure is similar to Figures 3B and 6A, and the three hydrogen bonds at the top of the figure are, from left to right, to Ala 165 O, Leu 164 N, and Asn 162 O. The lower hydrogen bond is to Tyr 223 N. **B:** Active site groove in the proenzyme. Hydrogen bonds between the residues 72–77 of the pro-peptide and the active site groove of prostromelysin are indicated by dashed lines. The orientation is similar to Figures 3A and 6B. The hydrogen bonding groups in the catalytic domain are the same as in Figure 3A with the addition of a hydrogen bond to Pro 221 O (lower left). Light blue sphere at the top center of the figure is calcium ion 1.

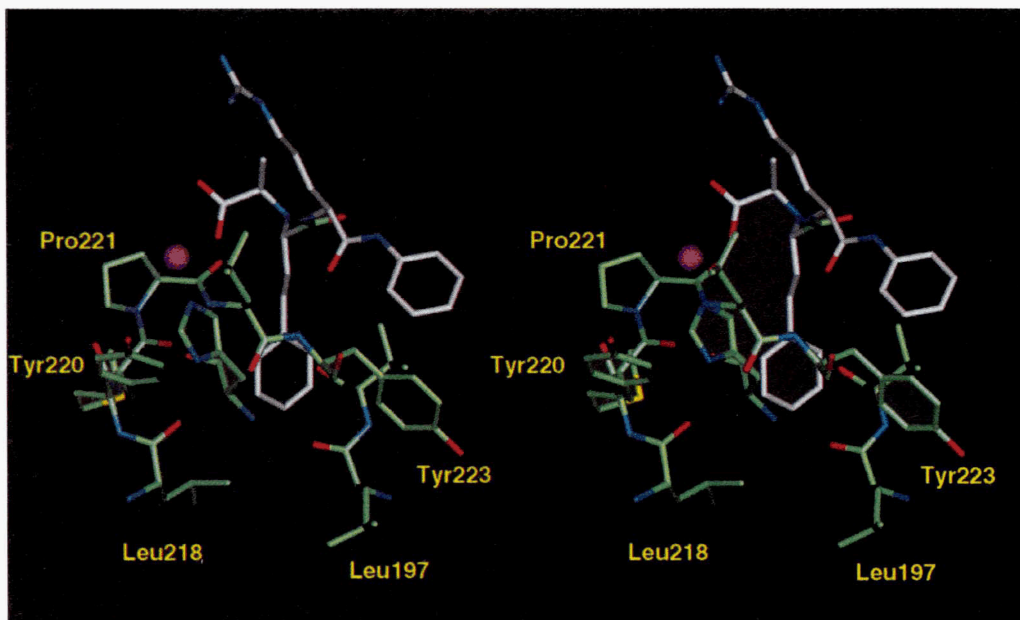

Fig. 4. The S1′ site. Inhibitor **I** (white carbon atoms) is shown together with all groups of the catalytic core that surround the homophenylalanyl group (light green carbon atoms). Catalytic zinc is represented as a violet sphere, and the zinc ligand His 201 is shown at the rear of the figure. Carbonyl groups of Leu 218 and Tyr 220 point at the edge of the aromatic ring of the inhibitor.

encirclement of this ring is completed by Leu 197–Val 198. At the *para*-position of the ring, the association is less intimate: the side chains of Leu 197 and Leu 218 flank this end of the molecule and form a gateway into the rest of S1′, a large hydrophobic tunnel that extends through the full span of the molecule (Fig. 5).

Discussion

Comparison of the hydrogen bonds formed between MMPs and their inhibitors reveals a general pattern of conservation with two notable variations (Fig. 6). All inhibitors and the propeptide of prostromelysin form β-structure-type hydrogen bonds

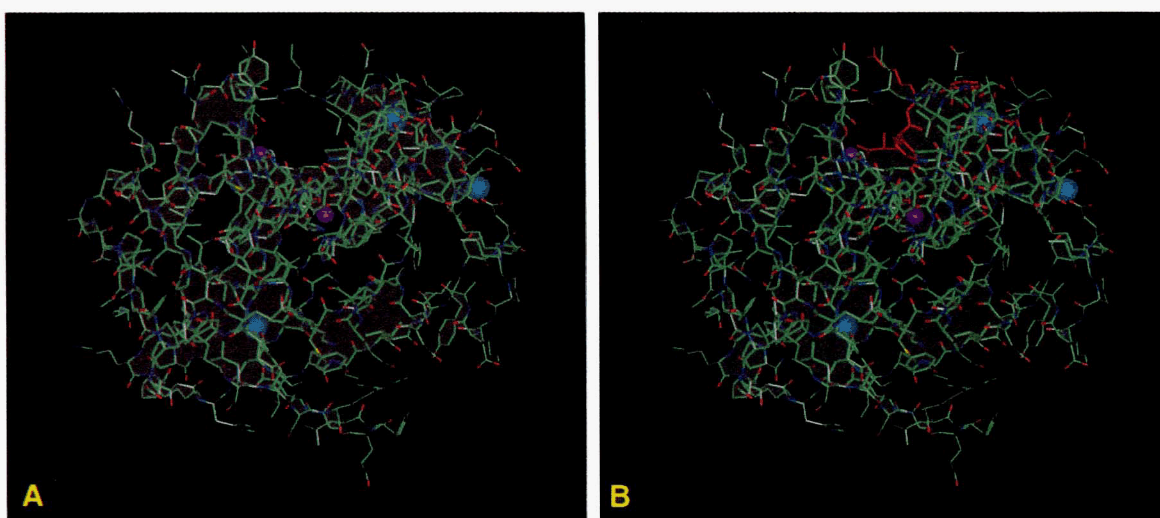

Fig. 5. Global view of the S1′ site. All-atom representation of the complex between stromelysin and **I**. Zinc and calcium ions are shown as violet and light blue spheres, respectively. Bound inhibitor (top center of figure) is drawn in red. The view is from the bottom of S1′, approximately from the bottom of Figure 3. **A:** Catalytic core and metal ions without the bound inhibitor. **B:** Identical to A including the bound inhibitor. Comparison of A and B indicates that the S1′ site extends through the body of the molecule and is open to solvent at both ends. This figure is not a stereo pair.

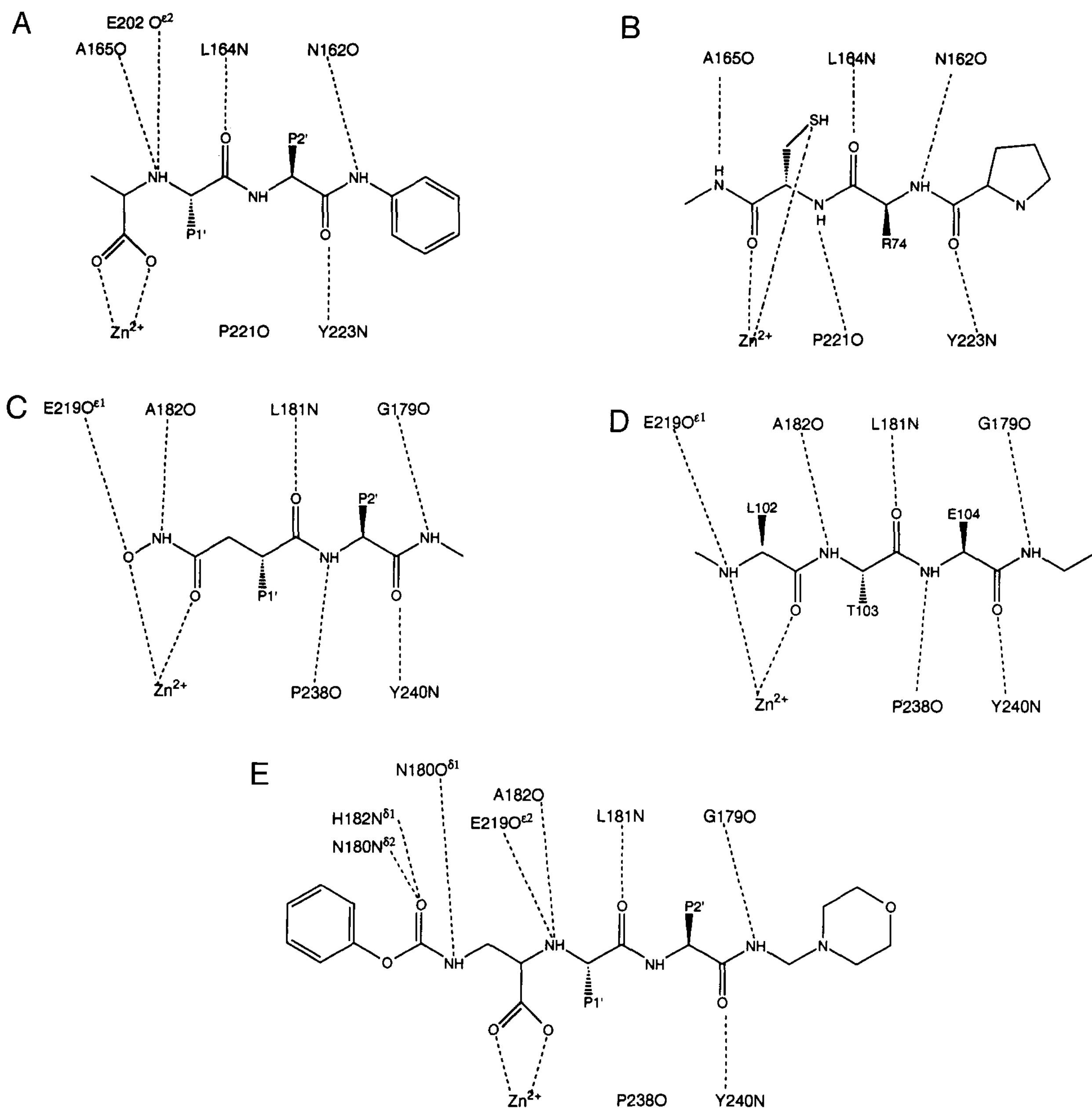

Fig. 6. Hydrogen bonds in the stromelysin active site. **A:** The complex with **I**. Compare with Figure 3A. **B:** The proenzyme propeptide. Compare with Figure 3B. **C:** Interaction between human fibroblast collagenase and a hydroxylamine inhibitor (Spurlino et al., 1994). Similar interactions, except that involving Glu 219$O^{\epsilon 1}$, were observed in another inhibited complex of human fibroblast collagenase (Borkakoti et al., 1994). Similar interactions were observed in an inhibited complex of human neutrophil collagenase (Stams et al., 1994). **D:** Interactions between human fibroblast collagenase and the N-terminus of an adjacent molecule in the crystal lattice (Lovejoy et al., 1994b). **E:** Interactions between human fibroblast collagenase and a carboxyalkyl amine inhibitor (Lovejoy et al., 1994a).

with residues that line the peptide-binding groove of the active site. It is particularly remarkable that the propeptide (Fig. 6B) forms hydrogen bonds of this type with the same groups of the catalytic core despite having an orientation opposite to that of all the other inhibitors. The inherent twofold symmetry of β-structures allows this groove to form a three-stranded β-sheet where the inhibitor is the central strand, but where that strand can adopt either chain direction. This observation is similar to the finding that SH3 domains can bind helical peptides in both orientations (Feng et al., 1994; Goudreau et al., 1994; Lim et al., 1994; Terasawa et al., 1994). Second, the hydrogen bond between the P2′ nitrogen atom and Pro 221 (or Pro 238 in the collagenases) is not formed in every case: prostromelysin and two forms of inhibited collagenase (Fig. 6B,C,D) have this hydrogen bond, but the stromelysin complex with **I** and another form of inhibited collagenase (Fig. 6A,E) do not, despite the fact that there is no large conformational change involving this residue. The opportunity to reverse the sense of the peptide chain

and to eliminate one peptide element of inhibitors presents a clear opportunity for the design of novel and specific inhibitors of these enzymes. In fact, a compound lacking the P2′ nitrogen has been described as an MMP inhibitor (MacPherson, 1994).

The structures reported here provide strong evidence for the cysteine-switch model of activation (Springman et al., 1990; Van Wart & Birkedal-Hansen, 1990; Salowe et al., 1992): the active site of the proenzyme is filled by the propeptide and Cys 75 interacts directly with the catalytic zinc ion, but the structure of the catalytic domain is relatively unchanged by activation (Fig. 7). Previous studies have suggested that several paths can lead to the activation of prostromelysin. Heat activation apparently depends on the presence of a small amount of active enzyme and results in cleavage at the His 82–Phe 83 bond to produce the mature enzyme (P.M. Cameron, 1995, in press). Activation by other proteases apparently proceeds through a stepwise mechanism, with early cleavage at sites involving residues 34–39 and subsequent cleavage after residue 83 (Nagase et al., 1991). Activation by mercurial reagents was originally believed to involve an initial cleavage of the Glu 68–Val 69 bond (Nagase et al., 1990), but recent studies have shown that cleavages after residues 16, 20, 24, 34, 50, 53, 57–59, 69–70, and 83 also occur (P.M. Cameron, 1995, in press). Of these sites (Figs. 7, 8), those involving residues 56–59 are in the loop between helices 2 and 3 in the proenzyme and appear readily accessible without any substantial conformational change in the structure of the proenzyme. Similarly, the bond between residues 15 and 16 and the bonds involving residues 33–39 are probably solvent accessible in the zymogen. The other cleavage sites are located in the three helices of the pro-domain, and it is likely that some conformational change, probably following an earlier cleavage, must precede reactions at these sites.

Activation also appears to involve a substantial rearrangement of residues 83–89 (Fig. 7). When the structures of the mature and proenzymes are aligned, residue 83 in one structure is more than 17 Å away from the corresponding residue in the other. By residue 90, however, the two structures are in register and remain remarkably similar throughout the remainder of the chain. In prostromelysin, residues 83–89 form part of a large loop that leads from helix A through the active site groove. During the activation process, these residues move to a completely different position, terminating with a salt link between the amine nitrogen of Phe 83 and the side chain of Asp 237 in helix C. This interaction is strikingly similar to that observed in "superactivated" collagenases. In that system, active forms of collagenase that have Phe 79 (homologous to Phe 83 in stromelysin) at their N-termini display 2–12-fold higher specific activity than those whose chains start at residues 80 or 81 (Murphy et al., 1987; Suzuki et al., 1990; Knäuper et al., 1993). In the three-dimensional structure of human neutrophil collagenase containing Phe 79 as the N-terminus, a salt link to the residue homologous to Asp 237, Asp 232, which is conserved among all MMPs, is formed (Suzuki et al., 1990; Reinemer et al., 1994). In the structures of forms of this protein with different N-termini, and correspondingly lower activities, the N-terminal residues are disordered, but the active sites of all of these enzymes are remarkably similar. These observations led to the hypothesis that disorder of the N-terminus in protein lacking Phe 79 interferes with substrate binding, and that the Phe 79–Asp 232 salt link prevents such interference (Reinemer et al., 1994). The conservation of this feature in the structure of mature stromelysin lends support to this hypothesis and suggests that the salt link from the N-terminus to the conserved Asp residue is a general feature of MMP structure.

This salt link is sufficiently distant from the catalytic residues (12 Å from the catalytic zinc ion, Fig. 2A) that its formation does not have a direct effect on the catalytic or substrate recognition sites. In addition, activation does not cause a substantial re-

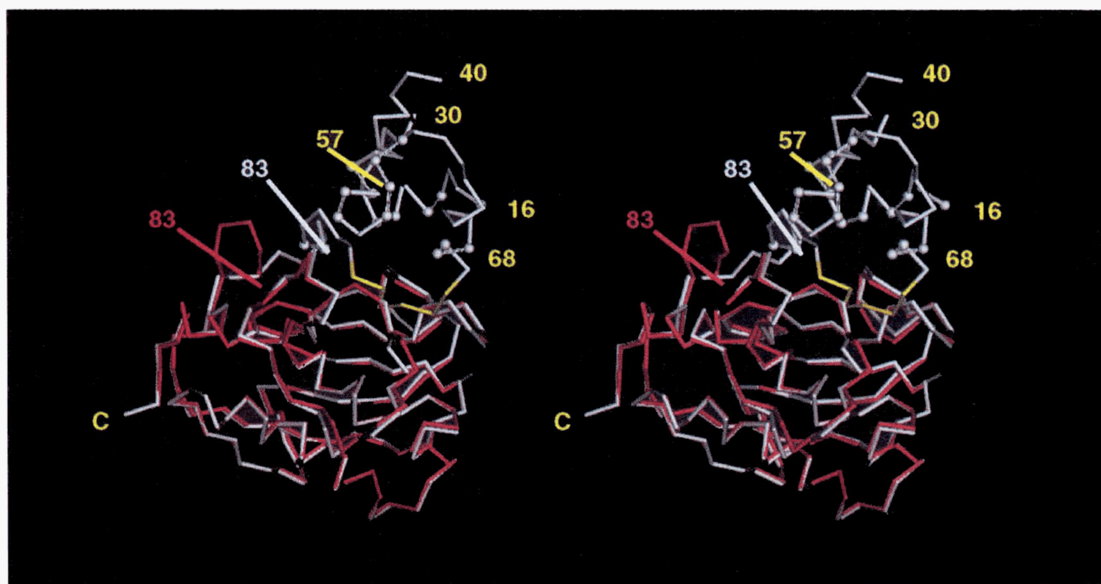

Fig. 7. Comparison of prostromelysin with the active enzyme. α-Carbon traces of prostromelysin (white) and active stromelysin (red) are superimposed. Residues 72–77 of the propeptide, which lie in the active site groove, are drawn in yellow. White spheres represent peptide bonds that have been reported to be cleaved during mercurial activation of stromelysin. Residue 83 moves more than 17 Å during activation.

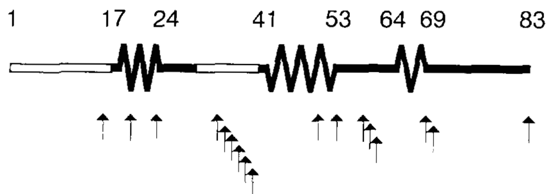

Fig. 8. Schematic representation of the pro-domain of prostromelysin. Regions belonging to α-helices are represented as jagged lines and regions with no regular secondary structures as thin bars. Parts of the chain that are not visible in the electron density maps are shown as open bars. Sequence numbers refer to the first and last residues of the prodomain and of the helices. Arrows indicate sites of cleavage during the activation of stromelysin.

arrangement of these sites. When the structures of the catalytic domain and proenzyme are aligned using all atoms of the residues that make contact with the bound inhibitor as guides, these atoms differ by an average of 0.37 Å. The largest difference is at Pro 221, where the average difference is 0.89 Å. This residue participates in a hydrogen bond with the propeptide (Fig. 6B) but does not make a corresponding interaction with the bound inhibitor (Fig. 6A). The carbonyl oxygen of this residue differs by 0.74 Å between the two structures.

Comparison of the amino acid sequences of the pro-domains of the MMP family suggests that they are likely to have similar structures. When the sequence of the prodomain of stromelysin-1 is aligned with those of stromelysin-2, matrilysin, 72-kDa and 92-kDa gelatinases, and fibroblast and neutrophil collagenases, it shows 38–77% identity with these other proteins (data not shown). A multiple alignment of these sequences shows two regions of above average similarity (Fig. 9): a region corresponding to the first helix and a second region encompassing the second and third helices and the peptide that lies in the active site. Two of the residues from the first helix (Tyr 20 and Leu 21) and residues surrounding Cys 75 are nearly invariant among MMPs, and mutagenesis studies have established their importance in maintaining the latent state (Sanchez-Lopez et al., 1988; Park et al., 1991; Freimark et al., 1994). These sequence similarities

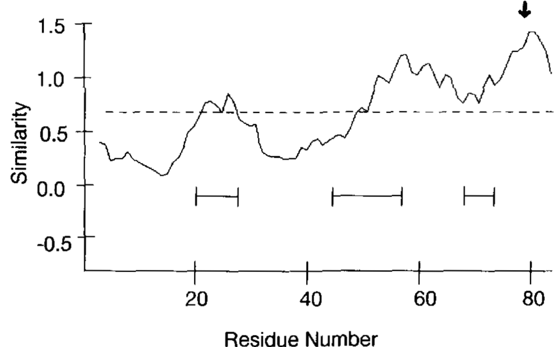

Fig. 9. Plot of similarity of amino acid sequences among the prodomains of MMPs calculated with the program PLOTSIMILARITY (Genetics Computer Group, Madison, Wisconsin). Horizontal bars correspond to the three helices observed in the structure of prostromelysin-255. Arrow indicates the position of the conserved cysteine residue that interacts with the catalytic zinc ion.

suggest that the prodomains of these proteins resemble one another in three-dimensional structure, as the catalytic domain of stromelysin-1 resembles the catalytic domains of fibroblast and neutrophil collagenase.

Comparison of the structures of stromelysin and thermolysin reveals that several residues proposed to play critical roles in the proteolytic mechanism of thermolysin (Matthews, 1988) are conserved in stromelysin (Figs. 10, 11). However, three residues proposed to play a stabilizing role in the thermolysin mechanism have no counterparts in stromelysin. The essential features of the thermolysin mechanism involve nucleophilic attack by activated water on the carbonyl group of the scissile bond to form a hemiketal. The carbonyl group is a ligand of the catalytic zinc ion and the water molecule is hydrogen bonded with the side chain of Glu 143, which appears to function as a general acid/base during catalysis. The side chains of Tyr 157 and His 231 stabilize the oxyanion of the hemiketal, whereas the carbonyl oxygen atom of Ala 113 and the side chain of Asn 112 hydrogen bond to the amide nitrogen of the scissile bond (Fig. 10A). Superimposition of the active sites of stromelysin and thermolysin reveal an equivalent, but smaller, set of catalytic residues in stromelysin (Fig. 11). The zinc ions of both proteins are positioned to play equivalent roles as are Glu 143 (thermolysin) and Glu 202 (stromelysin). Consistent with this model, substitution of the analogous glutamyl residue in gelatinase A with an aspartyl residue causes a 100-fold decrease in specific activity; mutants with alanine or glutamine at this site have only 0.01% of wild-type activity (Crabbe et al., 1994). Further comparison indicates that the carbonyl oxygen atom of Ala 165 in stromelysin appears to act analogously to the carbonyl group of Ala 113 of thermolysin. Surprisingly, the volumes occupied by the side chains of Tyr 157 and His 231 in thermolysin are empty in stromelysin, and there do not appear to be any nearby groups that can play an equivalent role in stabilizing the oxyanion of the hemiketal. The volume occupied by the side chain of Asn 112 in thermolysin is occupied by the side chain of Val 163 in stromelysin, a group clearly incapable of hydrogen bonding with the scissile amide nitrogen. A similar lack of thermolysin-equivalent residues has been reported in the structure of C-truncated human fibroblast collagenase (Lovejoy et al., 1994b; Spurlino et al., 1994). In one of these studies, it was suggested that the side chain of Asn 180 (collagenase) may substitute for the side chain of Asn 112 (thermolysin) (Spurlino et al., 1994). Such a substitution is not possible in the structure presented here. It is possible that in full-length stromelysin or collagenase the sites equivalent to the thermolysin residues Asn 112, Tyr 157, and His 231 may be occupied by groups from the C-terminal domain. However, the observation (Marcy et al., 1991; Ho et al., 1994) that removal of this domain does not significantly affect the peptidase activity, specificity, and inhibition of these enzymes makes it unlikely that such residues play a significant role in catalysis. The present comparison of stromelysin and thermolysin suggests that the zinc ion, glutamyl side chain, and carbonyl groups that are shared by these proteins constitute a minimal constellation of catalytic elements necessary to support proteolysis by this class of enzymes.

The active site of stromelysin is characterized by a catalytic zinc ion that lies in a groove containing several protein groups capable of making β-structure-like hydrogen bonds with substrates and inhibitors as well as an extraordinarily large hydrophobic S1′ recognition site. Comparison of the structures presented here

Fig. 10. Thermolysin and stromelysin catalytic residues and mechanism. **A:** Proposed mechanism for thermolysin (Matthews, 1988). **B:** Analogous mechanism for proteolysis by stromelysin. Three residues (Tyr 157, His 231, and Asn 112) believed to participate in thermolysin-catalyzed proteolysis have no equivalents in stromelysin.

Fig. 11. Superposition of the active sites of thermolysin (thin bonds, PDB [Bernstein et al., 1977] entry 4TMN) and stromelysin (thick bonds). The zinc ion of thermolysin is drawn as a sphere. The thermolysin structure also contains the transition state analogue carbobenzoxy-L-PheP-L-Leu-L-Ala. (PheP indicates that the trigonal carbon of the peptide linkage is replaced by the tetrahedral phosphorus of a phosphonamidate group.) Three residues in thermolysin (white captions), Asn 112, Tyr 157, and His 231, do not have structural equivalents in stromelysin (yellow captions).

Table 1. *Data collection and reduction*

Structure	Detector	d_{min}[a] (Å)	Total observations	Unique reflections	R_m[b] (%)	Completeness (%)
SLN255:**I**	Siemens	2.26	22,355	8,044	7.58	83.1
proSLN255	R-AXIS IIC	1.90	34.939	24,166	5.55	98.2

[a] d_{min} is the Bragg spacing of the highest resolution datum used in the structure solution and refinement.
[b] R_m, the R-factor corresponding to the averaging of equivalent reflections during data reduction. $R_m = \sum[|I(h) - I_{av}(h)|]/\sum I_{av}(h)$, where $I(h)$ is an individual measurement of the intensity of reflection h; and $I_{av}(h)$ is the average of multiple observations of reflection h. The summation is over all reflections h.

with the previously known zinc metalloprotease structures extends our understanding of the essential features of substrate recognition and catalysis by these enzymes. The large S1′ site of stromelysin, in particular, contains an extensive volume that is not even half filled by the homophenylalanyl group of the inhibitor used in this study. The exploitation of this large hydrophobic space, punctuated by hydrophilic groups, particularly carbonyl oxygens directed toward the hydrophobic P1′ side chain, represents an opportunity for the structure-based design of more potent and more specific inhibitors of this enzyme. The fact that the peptide chain of inhibitors can interact with the active site groove in two different orientations presents an additional opportunity to design compounds of particular potency and selectivity.

Materials and methods

Recombinant human proSLN-255 was expressed in *Escherichia coli*, purified, and activated to the mature form (residues 83–255) by published procedures (Marcy et al., 1991; Gooley et al., 1993). Both proSLN-255 and the inhibited complex of mature SLN-255 were crystallized by hanging drop vapor diffusion. For proSLN-255, protein solution (10 mg/mL proSLN-255, 5 mM $CaCl_2$, 0.05 mM $Zn[OAc]_2$, 0.02% NaN_3, 20 mM MES, pH 6.5) was mixed with an equal volume of reservoir buffer (14% PEG-6000, 5% sat. Na-citrate, 100 mM cacodylate, pH 5.8) and incubated at 4 °C. The protein solution also contained a trace amount of an inhibitor to prevent autoactivation (approximately 0.02 mole inhibitor per mole protein). Crystals formed after several weeks and belong to the orthorhombic space group F222, with $a = 111.04$, $b = 145.56$, $c = 76.75$ Å. For the complex, the protein solution contained 9 mg/mL SLN-255, 1.5 mM **I**, 5.0 mM $CaCl_2$, 0.02% NaN_3, 20 mM Tris-HCl, pH 7.5, and the reservoir contained 10% PEG-6000, 15% sat. NH_4OAc, 0.02% NaN_3, 0.1 M cacodylate, pH 5.54. The crystals belong to the trigonal space group $P3_121$, with $a = 47.23$, $c = 150.85$ Å. Three-dimensional diffraction data (Table 1) were collected using a Siemens area detector (inhibited complex) and an R-Axis IIC area detector (proSLN-255) and CuKα radiation from a Rigaku RU-200 rotating anode X-ray generator. Data were processed using the XENGEN (Howard et al., 1987), RAXIS (R-Axis IIc Data Processing Software v. 2.1, Rigaku Corporation, Tokyo), and FBSCALE (Weissman, 1982) packages.

Both structures were solved by molecular replacement methods using X-PLOR (Brünger, 1990) and a probe consisting of the protein portion of the stromelysin complex with a different inhibitor, which had been solved by heavy-atom methods (J.W. Becker et al., in prep). Electron density corresponding to the portions of these structures that are not in the search probe (e.g., the pro-peptide and the bound inhibitor) was clearly visible in initial maps. Complete models were constructed by interactive model-building (Sack, 1988) and refinement using X-PLOR including one cycle of simulated annealing (Brünger, 1992b). A bulk solvent mask was included in the model, and in the later stages of each refinement, each model was confirmed by 10% simulated-annealing omit maps (Hodel et al., 1992). The final model of the inhibited complex comprises residues 83–250, 2 Zn^{2+} ions, 3 Ca^{2+} ions, the bound inhibitor, and 51 ordered water molecules. For proSLN-255, the model contains residues 16–30, 41–250, 2 Zn^{2+} ions, 2 Ca^{2+} ions, and 119 ordered water molecules. Final refinement parameters are presented in Table 2. At least some of the missing portions of the proSLN-255 chain may be due to proteolysis. Electrospray mass spectral analysis indicates that freshly prepared protein is homogeneous and of the expected mass, but that protein from redissolved crystals contains several species (data not shown).

Protein sequences were analyzed using the GCG package (Genetics Computer Group, Madison, Wisconsin).

Acknowledgments

We thank Patrick R. Griffin and Tracey D. Klatt for the mass spectral analysis of stromelysin, and Nancy Thornberry, Paul Gooley, and Scott Salowe for critical reading of the manuscript. The coordinates of both structures will be deposited in the Protein Data Bank.

Table 2. *Refinement statistics*

Name	R[a]	R_{free}[b]	⟨δ(Bonds)⟩[c] (Å)	⟨δ(Angles)⟩[c]
SLN255:**I**	0.2232	0.2975	0.009	1.440°
proSLN255	0.2185	0.2569	0.009	1.329°

[a] R, the crystallographic R-factor for the refined structures. $R = \sum[|F(h) - F_{av}(h)|]/\sum I_{av}(h)$.
[b] R_{free}, the cross-validation R-factor (Brünger, 1992a) corresponding to 10% of the reflection data that were omitted from refinement of the structures.
[c] RMS deviations of bond lengths and angles in the refined structures from ideal values (Engh & Huber, 1991).

References

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535-542.

Birkedal-Hansen H, Moore WGI, Bodden MK, Windsor LJ, Birkedal-Hansen B, DeCarlo A, Engler JA. 1993. Matrix metalloproteinases: A review. *Crit Rev Oral Biol Med 4*:197-250.

Bode W, Gomis-Rüth FX, Stockler W. 1993. Astacins, serralysins, snake-venom and matrix metalloproteinases exhibit identical zinc-binding environments (hexxhxxgxxh and met-turn) and topologies and should be grouped into a common family, the metzincins. *FEBS Lett 331*:134-140.

Bode W, Reinemer P, Huber R, Kleine T, Schnierer S, Tschesche H. 1994. The X-ray crystal-structure of the catalytic domain of human neutrophil collagenase inhibited by a substrate-analog reveals the essentials for catalysis and specificity. *EMBO J 13*:1263-1269.

Borkakoti N, Winkler FK, Williams DH, Darcy A, Broadhurst MJ, Brown PA, Johnson WH, Murray EJ. 1994. Structure of the catalytic domain of human fibroblast collagenase complexed with an inhibitor. *Nature Struct Biol 1*:106-110.

Brünger AT. 1990. Extension of molecular replacement: A new search strategy based on Patterson correlation refinement. *Acta Crystallogr A 46*: 46-57.

Brünger AT. 1992a. Free R value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature 355*:472-475.

Brünger AT. 1992b. *X-PLOR: Version 3.1, a system for X-ray crystallography and NMR*. New Haven/London: Yale University Press.

Cameron PM, Marcy AI, Robosz LL, Hermes JD. 1995. Use of an active site inhibitor to elucidate the mechanism of prostromelysin activation. *Bioorganic Chemistry* in press.

Chapman KT, Kopka IE, Durette PL, Esser CK, Lanza TJ, Izquierdo-Martin M, Niedzwiecki L, Chang B, Harrison RK, Kuo DW, Lin TY, Stein RL, Hagmann WK. 1993. Inhibition of matrix metalloproteinases by *N*-carboxyalkyl peptides. *J Med Chem 36*:4293-4301.

Chin JR, Murphy G, Werb Z. 1985. Stromelysin, a connective tissue-degrading metalloendopeptidase secreted by stimulated rabbit synovial fibroblasts in parallel with collagenase. Biosynthesis, isolation, characterization, and substrates. *J Biol Chem 260*:12367-12376.

Crabbe T, Zucker S, Cockett MI, Willenbrock F, Tickle S, O'Connell JP, Scothern JM, Murphy G, Docherty A. 1994. Mutation of the active-site glutamic-acid of human gelatinase-a – Effects on latency, catalysis, and the binding of tissue inhibitor of metalloproteinases-1. *Biochemistry 33*:6684-6690.

Docherty AJP, O'Connell J, Crabbe T, Angal S, Murphy G. 1992. The matrix metalloproteinases and their natural inhibitors – Prospects for treating degenerative tissue diseases. *Trends Biotechnol 10*:200-207.

Engh RA, Huber R. 1991. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr A 47*:392-400.

Feng SB, Chen JK, Yu HT, Simon JA, Schreiber SL. 1994. Two binding orientations for peptides to the src SH3 domain – Development of a general-model for SH3-ligand interactions. *Science 266*:1241-1247.

Freimark BD, Feeser WS, Rosenfeld SA. 1994. Multiple sites of the propeptide region of human stromelysin-1 are required for maintaining a latent form of the enzyme. *J Biol Chem 269*:26982-26987.

Galloway WA, Murphy G, Sandy JD, Gavrilovic J, Cawston TE, Reynolds JJ. 1983. Purification and characterization of a rabbit bone metalloproteinase that degrades proteoglycan and other connective-tissue components. *Biochem J 209*:741-752.

Gomis-Rüth FX, Kress LF, Kellermann J, Mayr I, Lee X, Huber R, Bode W. 1994. Refined 2 Å X-ray crystal-structure of the snake-venom zinc-endopeptidase adamalysin. II. Primary and tertiary structure determination, refinement, molecular-structure and comparison with astacin, collagenase and thermolysin. *J Mol Biol 239*:513-544.

Gooley PR, Johnson BA, Marcy AI, Cuca GC, Salowe SP, Hagmann WK, Esser CK, Springer JP. 1993. Secondary structure and zinc ligation of human recombinant short-form stromelysin by multidimensional heteronuclear NMR. *Biochemistry 32*:13098-13108.

Gooley PR, O'Connell JF, Marcy AI, Cuca GC, Salowe SP, Bush BL, Hermes JD, Esser CK, Hagmann WK, Springer JP, Johnson BA. 1994. The NMR structure of the inhibited catalytic domain of human stromelysin-1. *Nature Struct Biol 1*:111-118.

Goudreau N, Cornille F, Duchesne M, Parker F, Tocqué B, Garbay C, Roques BP. 1994. NMR structure of the N-terminal SH3 domain of GRB2 and its complex with a proline-rich peptide from Sos. *Nature Struct Biol 1*:898-907.

Gowen M, Wood DD, Ihrie EJ, Meats JE, Russell RG. 1984. Stimulation by human interleukin 1 of cartilage breakdown and production of collagenase and proteoglycanase by human chondrocytes but not by human osteoblasts in vitro. *Biochim Biophys Acta 797*:186-193.

He C, Wilhelm SM, Pentland AP, Marmer BL, Grant GA, Eisen AZ, Goldberg GI. 1989. Tissue cooperation in a proteolytic cascade activating human interstitial collagenase. *Proc Natl Acad Sci USA 86*:2632-2636.

Ho TF, Qoronfleh MW, Wahl RC, Pulvino TA, Vavra KJ, Falvo J, Banks TM, Brake PG, Ciccarelli RB. 1994. Gene expression, purification and characterization of recombinant human neutrophil collagenase. *Gene 146*:297-301.

Hodel A, Kim SH, Brünger AT. 1992. Model bias in macromolecular crystal structures. *Acta Crystallogr A 48*:851-858.

Howard AJ, Gilliland GL, Finzel BC, Poulos TL, Ohlendorf DH, Salemme FR. 1987. The use of an imaging proportional counter in macromolecular crystallography. *J Appl Crystallogr 20*:383-387.

Knäuper V, Wilhelm SM, Seperack PK, DeClerck YA, Langley KE, Osthues A, Tschesche H. 1993. Direct activation of human neutrophil procollagenase by recombinant stromelysin. *Biochem J 295*:581-586.

Koklitis PA, Murphy G, Sutton C, Angal S. 1991. Purification of recombinant human prostromelysin. Studies on heat activation to give high-Mr and low-Mr active forms, and a comparison of recombinant with natural stromelysin activities. *Biochem J 276*:217-221.

Lim WA, Richards FM, Fox RO. 1994. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature 372*:375-379.

Lovejoy B, Cleasby A, Hassell AM, Longley K, Luther MA, Weigl D, McGeehan G, McElroy AB, Drewry D, Lambert MH, Jordan SR. 1994a. Structure of the catalytic domain of fibroblast collagenase complexed with an inhibitor. *Science 263*:375-377.

Lovejoy B, Hassell AM, Luther MA, Weigl D, Jordan SR. 1994b. Crystal-structures of recombinant 19-kda human fibroblast collagenase complexed to itself. *Biochemistry 33*:8207-8217.

MacNaul KL, Chartrain N, Lark M, Tocci MJ, Hutchinson NI. 1990. Discoordinate expression of stromelysin, collagenase, and tissue inhibitor of metalloproteinases-1 in rheumatoid human synovial fibroblasts. Synergistic effects of interleukin-1 and tumor necrosis factor-alpha on stromelysin expression. *J Biol Chem 265*:17238-17245.

MacPherson LJ. 1994. Arylsulfonamido-substituted hydroxamic acids. European Patent Application 0-606-046-A1.

Marcy AI, Eiberger LL, Harrison R, Chan HK, Hutchinson NI, Hagmann WK, Cameron PM, Boulton DA, Hermes JD. 1991. Human fibroblast stromelysin catalytic domain: Expression, purification, and characterization of a C-terminally truncated form. *Biochemistry 30*:6476-6483.

Matthews BH. 1988. Structural basis of the action of thermolysin and related zinc peptidases. *Acc Chem Res 21*:333-340.

Murphy G, Cockett MI, Stephens PE, Smith BJ, Docherty AJP. 1987. Stromelysin is an activator of procollagenase. *Biochem J 248*:265-268.

Murphy G, McAlpine CG, Poll CT, Reynolds JJ. 1985. Purification and characterization of a bone metalloproteinase that degrades gelatin and types IV and V collagen. *Biochim Biophys Acta 831*:49-58.

Murphy GJP, Murphy G, Reynolds JJ. 1991. The origin of matrix metalloproteinases and their familial relationships. *FEBS Lett 289*:4-7.

Murrell GAC, Jang D, Williams RJ. 1995. Nitric oxide activates metalloprotease enzymes in articular cartilage. *Biochem Biophys Res Commun 206*:15-21.

Nagase H, Enghild JJ, Suzuki K, Salvesen G. 1990. Stepwise activation mechanisms of the precursor of matrix metalloproteinase 3 (stromelysin) by proteinases and (4-aminophenyl)mercuric acetate. *Biochemistry 29*:5783-5789.

Nagase H, Suzuki K, Enghild JJ, Salvesen G. 1991. Stepwise activation mechanisms of the precursors of matrix metalloproteinases 1 (tissue collagenase) and 3 (stromelysin). *Biomed Biochim Acta 50*:749-754.

Ogata Y, Enghild JJ, Nagase H. 1992. Matrix metalloproteinase 3 (stromelysin) activates the precursor for the human matrix metalloproteinase 9. *J Biol Chem 267*:3581-3584.

Okada Y, Harris ED Jr, Nagase H. 1988. The precursor of a metalloendopeptidase from human rheumatoid synovial fibroblasts. Purification and mechanisms of activation by endopeptidases and 4-aminophenylmercuric acetate. *Biochem J 254*:731-741.

Okada Y, Konomi H, Yada T, Kimata K, Nagase H. 1989. Degradation of type IX collagen by matrix metalloproteinase 3 (stromelysin) from human rheumatoid synovial cells. *FEBS Lett 244*:473-476.

Okada Y, Nagase H, Harris ED Jr. 1986. A metalloproteinase from human rheumatoid synovial fibroblasts that digests connective tissue matrix components. Purification and characterization. *J Biol Chem 261*:14245-14255.

Okada Y, Nakanishi I. 1989. Activation of matrix metalloproteinase 3 (stromelysin) and matrix metalloproteinase 2 ("gelatinase") by human neutrophil elastase and cathepsin G. *FEBS Lett 249*:353-356.

Park AJ, Matrisian LM, Kells AF, Pearson R, Yuan ZY, Navre M. 1991. Mutational analysis of the transin (rat stromelysin) autoinhibitor region demonstrates a role for residues surrounding the "cysteine switch." *J Biol Chem 266*:1584–1590.

Reinemer P, Grams F, Huber R, Kleine T, Schnierer S, Piper M, Tschesche H, Bode W. 1994. Structural implications for the role of the N-terminus in the superactivation of collagenases — A crystallographic study. *FEBS Lett 338*:227–233.

Sack JS. 1988. CHAIN — A crystallographic modeling program. *J Mol Graphics 6*:224–225.

Salowe SP, Marcy AI, Cuca GC, Smith CK, Kopka IE, Hagmann WK, Hermes JD. 1992. Characterization of zinc-binding sites in human stromelysin-1: Stoichiometry of the catalytic domain and identification of a cysteine ligand in the proenzyme. *Biochemistry 31*:4535–4540.

Sanchez-Lopez R, Nicholson R, Gesnel MC, Matrisian LM, Breathnach R. 1988. Structure-function relationships in the collagenase family member transin. *J Biol Chem 263*:11892–11899.

Saus J, Quinones S, Otani Y, Nagase H, Harris ED Jr, Kurkinen M. 1988. The complete primary structure of human matrix metalloproteinase-3. Identity with stromelysin. *J Biol Chem 263*:6742–6745.

Schmid TM, Mayne R, Jeffrey JJ, Linsenmayer TF. 1986. Type X collagen contains two cleavage sites for a vertebrate collagenase. *J Biol Chem 261*:4184–4189.

Seltzer JL, Eisen AZ, Bauer EA, Morris NP, Glanville RW, Burgeson RE. 1989. Cleavage of type VII collagen by interstitial collagenase and type IV collagenase (gelatinase) derived from human skin. *J Biol Chem 264*: 3822–3826.

Springman EB, Angleton EL, Birkedal-Hansen H, Van Wart HE. 1990. 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 USA 87*:364–368.

Spurlino JC, Smallwood AM, Carlton DD, Banks TM, Vavra KJ, Johnson JS, Cook ER, Falvo J, Wahl RC, Pulvino TA, Wendoloski JJ, Smith DL. 1994. 1.56 Å structure of mature truncated human fibroblast collagenase. *Proteins Struct Funct Genet 19*:98–109.

Stams T, Spurlino JC, Smith DL, Wahl RC, Ho TF, Qoronfleh MW, Banks TM, Rubin B. 1994. Structure of human neutrophil collagenase reveals large S1′ specificity pocket. *Nature Struct Biol 1*:119–123.

Stöcker W, Grams F, Baumann U, Reinemer P, Gomis-Rüth FX, McKay DB, Bode W. 1995. The metzincins — Topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases. *Protein Sci 4*:823–840.

Suzuki K, Enghild JJ, Morodomi T, Salvesen G, Nagase H. 1990. Mechanisms of activation of tissue procollagenase by matrix metalloproteinase 3 (stromelysin). *Biochemistry 29*:10261–10270.

Terasawa H, Kohda D, Hatanaka H, Tsuchiya S, Ogura K, Natgata K, Ishii S, Mandiyan V, Ullrich A, Schlessinger J, Inagaki F. 1994. Structure of the N-terminal SH3 domain of GRB2 complexed with a peptide from the guanine nucleotide releasing factor Sos. *Nature Struct Biol 1*:891–897.

Van Wart HE, Birkedal-Hansen H. 1990. The cysteine switch: A principle of regulation of metalloproteinase activity with potential applicability to the entire matrix metalloproteinase gene family. *Proc Natl Acad Sci USA 87*:5578–5582.

Weissman L. 1982. Strategies for extracting isomorphous and anomalous signals. In: D. Sayre, ed. *Computational crystallography*. Oxford: Oxford University Press. pp 56–63.

Welgus HG, Jeffrey JJ, Eisen AZ. 1981. The collagen substrate specificity of human skin fibroblast collagenase. *J Biol Chem 256*:9511–9515.

Woessner JF Jr. 1991. Matrix metalloproteinases and their inhibitors in connective tissue remodeling. *FASEB J 5*:2145–2154.

Ye QZ, Johnson LL, Nordan I, Hupe D, Hupe L. 1994. A recombinant human stromelysin catalytic domain identifying tryptophan derivatives as human stromelysin inhibitors. *J Med Chem 37*:206–209.