Human Fibroblast Stromelysin Catalytic Domain: Expression, Purification, and Characterization of a C-Terminally Truncated Form

Alice I. Marcy,[‡] Laura L. Eiberger,[‡] Richard Harrison,[§] H. Karen Chan,[‡] Nancy I. Hutchinson,[∥] William K. Hagmann,[⊥] Patricia M. Cameron,[‡] David A. Boulton,[‡] and Jeffrey D. Hermes*[,‡]

Departments of Biophysical Chemistry, Enzymology, Molecular Immunology, and Medicinal Chemical Research, Merck Sharp & Dohme Research Laboratories, Rahway, New Jersey 07065

Received February 5, 1991; Revised Manuscript Received April 16, 1991

ABSTRACT: Stromelysin-1 is a member of a tissue metalloproteinase family whose members are all capable of degrading extracellular matrix components. A truncated form of human fibroblast prostromelysin 1 lacking the C-terminal, hemopexin-like domain has been expressed in *Escherichia coli* and purified to homogeneity. Treatment of this short form of prostromelysin with (aminophenyl)mercuric acetate resulted in activation and loss of the propeptide in a manner identical with the wild-type, full-length protein. Kinetic comparisons using Nle[11]-substance P as a substrate showed that the wild-type stromelysin and the truncated form of the enzyme had similar k_{cat} and K_m values. Likewise, both enzymes displayed similar K_i values for a hydroxamate-containing peptide inhibitor. Taken together, these results indicate that the C-terminal portion of stromelysin is not required for proper folding of the catalytic domain, maintenance of the enzyme in a latent form, activation with an organomercurial, cleavage of a peptide substrate, or interaction with an inhibitor. Moreover, the active short form of stromelysin displayed a reduction in the C-terminal heterogeneity, a characteristic degradation of the full-length stromelysin, and thereby provides a more suitable protein for future structural studies.

Stromelysin 1 (MMP-3)[1] is a member of the family of matrix metalloendoproteinases thought to be directly responsible for the proteolytic breakdown of cartilage. This proteolysis is postulated to result in a degradative loss of articular cartilage and the eventual impairment of joint function associated with both osteoarthritis (OA) and rheumatoid arthritis (RA). Members of this family are characterized by their high primary sequence homology, a catalytic Zn^{2+} requirement, inhibition by tissue inhibitors of metalloproteases (TIMP), and secretion as inactive proenzymes [reviewed by Matrisian (1990)]. Despite these biochemical similarities, the three best characterized members display quite distinct substrate specificities: tissue collagenase (MMP-1) cleaves collagen types I, II, III, VII, and X (Welgus et al., 1981; Schmid et al., 1986; Seltzer et al., 1989); gelatinase (MMP-2) degrades collagen types IV, V, and VII as well as gelatins (Seltzer et al., 1989; Murphy et al., 1985), stromelysin (MMP-3), showing an even broader specificity profile, cleaves proteoglycans, fibronectin, procollagen type I, collagen types III, IV, and IX, and laminin (Galloway et al., 1983; Chin et al., 1985; Okada et al., 1986; Okada & Nakanishi, 1989; Wilhelm et al., 1987, Quantin et al., 1989), and is implicated as a natural activator of collagenase (Murphy et al., 1987; Ito & Nagase, 1988; He et al., 1989; Brinckerhoff et al., 1990). Though the relative importance of each enzyme in the pathology associated with OA and RA has yet to be established, stromelysin, because of its ability to directly or indirectly degrade most of the major components of cartilage and its induction by the inflammatory mediator interleukin 1 (Gowen et al., 1984; MacNaul et al., 1990; Saus et al., 1988), appears to play a major role. Accordingly, the development of specific inhibitors of stromelysin could provide the next major class of therapeutics for the treatment of arthritis.

Stromelysin is secreted from cells as a 57-kDa monomeric proenzyme. Though a small fraction is secreted as an N-glycosylated form of 59 kDa (Wilhelm et al., 1987), this addition of carbohydrate appears to be unimportant for activity (Okada et al., 1988). The mechanism of activation of the proenzyme has been elucidated by use of kinetics and amino acid sequence analysis of activation intermediates and products (Nagase et al., 1990). These studies indicate that the removal of the N-terminal prosequence of approximately 80 amino acids, to yield the 45-kDa mature enzyme, occurs by various pathways of stepwise cleavages. The individual steps of this activation are either bimolecular with other proteinases or autolytically unimolecular or bimolecular. A highly conserved region of eight residues near the carboxy end of the prosequence, common to all the sequenced matrix metalloproteinases, contains a cysteine residue which presumably functions as a ligand of the proenzyme's active-site zinc. It has been proposed that activation of the proenzyme occurs when this sulfhydryl ligand is replaced with a water molecule (Van Wart & Birkedal-Hansen, 1990). Sulfhydryl group modifiers such as APMA may function as activators by virtue of their ability to tightly complex the cysteine sulfhydryl, thus exposing a zinc coordination site for the catalytic water molecule. After activation, the 45-kDa form of stromelysin further cleaves itself in an imprecise fashion to yield a set of active, but C-terminally heterogeneous forms of stromelysin of approximately 28 kDa (Okada et al., 1986, 1988, 1989;

* To whom correspondence should be addressed at R80M-203, Merck and Co., Inc., P.O. Box 2000, Rahway, NJ 07065.
[‡] Department of Biophysical Chemistry.
[§] Department of Enzymology.
[∥] Department of Molecular Immunology.
[⊥] Department of Medicinal Chemical Research.

[1] Abbreviations: MMP, matrix metalloproteinase; TIMP, tissue inhibitor of metalloproteinase; APMA, (4-aminophenyl)mercuric acetate; NMR, nuclear magnetic resonance; OA, osteoarthritis; RA, rheumatoid arthritis; sfSTR, short form of stromelysin; pro-sfSTR, short form of prostromelysin; PTH, phenylthiohydantoin; IPTG, isopropyl β-thiogalactoside; Nle-SP, norleucine[11]-substance P; HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid; SDS, sodium dodecyl sulfate; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; PUMP, putative metalloproteinase.

0006-2960/91/0430-6476$02.50/0

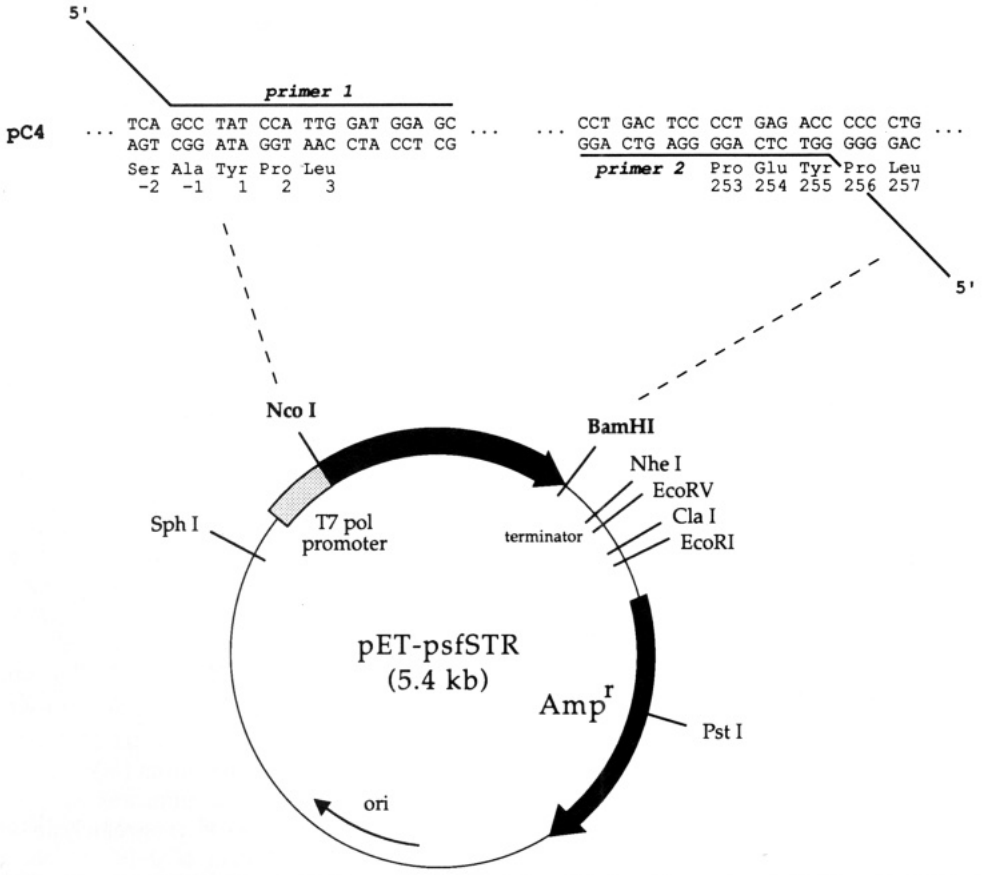


FIGURE 1: *E. coli* expression plasmid for the pro-sfSTR gene fragment. The pC4 plasmid contains the full-length cDNA for human fibroblast prostromelysin. Shown explicitly is the portion of the prostromelysin sequence which hybridizes to the PCR primers 1 and 2 (see text for detailed description of the primers). Ala_{-1} is the C-terminal residue of the natural, mammalian signal sequence for full-length prostromelysin. The expression plasmid pET-psfSTR was constructed by ligation of a PCR-generated fragment from pC4 into the *Nco*I and *Bam*HI sites of pET-8c (Studier et al., 1990).

Murphy et al., 1987; Nagase et al., 1990). The physiological relevance of this autolysis is unknown.

The observation that the 28-kDa forms of stromelysin are active led us to test whether direct expression of a C-terminally truncated form of prostromelysin would result in a latent proenzyme which could be activated to a C-terminally defined, stromelysin catalytic domain. This report describes the expression, purification, and characterization of a truncated form of human fibroblast stromelysin from *Escherichia coli*.

Materials and Methods

Materials. Oligonucleotides were synthesized on a Milligen 7500 DNA synthesizer and purified by using NENSORB PREP columns (NEN Research Products, Wilmington, DE) as described by Johnson et al. (1990). Salts for M9 medium and isopropyl β-D-thiogalactoside (IPTG) were obtained from Gibco BRL (Gaithersburg, MD); casamino acids were from Difco (Detroit, MI). Calcium chloride, sodium azide, sodium chloride, Tris-HCl, MES, zinc acetate, (aminophenyl)mercuric acetate (APMA), and soybean trypsin inhibitor attached to 4% beaded agarose were obtained from Sigma (St. Louis, MO). Enzyme-grade ammonium sulfate was obtained from Schwarz/Mann Biotech (ICN, Costa Mesa, CA). Brij-35 was purchased from Pierce (Rockford, IL). The chromatography resins Q-Sepharose, S-Sepharose, and Sephacryl S-100 HR were purchased from Pharmacia Fine Chemicals (Piscataway, NJ). Aprotinin, pepstatin, and leupeptin were from Boehringer Mannheim Biochemicals (Indianapolis, IN). Sodium dodecyl sulfate (SDS), glycine, acrylamide, protein molecular weight standards, and Coomassie Brilliant Blue R-250 were from Bio-Rad (Richmond, CA). Nor-Leu[11]-substance P (Nle-SP, Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-$NleNH_2$) was purchased from Bachem Bioscience (Philadelphia, PA). Recombinant human fibroblast prostromelysin expressed from mammalian cells was purchased from Celltech (Slough, U.K.). PCR reagents, including *Taq* polymerase, were obtained from Perkin Elmer Cetus (Norwalk, CT).

Expression Plasmid Construction. Plasmid pC4 contains the protein-coding region of human stromelysin 1 cDNA cloned from human synovial fibroblasts (MacNaul et al., 1990). To subclone a truncated version of stromelysin, two synthetic oligonucleotide primers were used with the pC4 cDNA template in a polymerase chain reaction (PCR) (Figure 1). Primer 1 (5′-ATACCATGGCCTATCCATTGG-ATGGAGC-3′) incorporates sequences for a unique *Nco*I site and an initiating methionine. Primer 2 (5′-ATAGGATCCTTAGGTCTCAGGGGAGTCAGG-3′) introduces sequences for a stop codon and a unique *Bam*HI site. The engineering of the *Nco*I site (CCATGG) required that the first nucleotide after the initiating methionine (ATG) be a G. For this reason, Ala_{-1} of the preprostromelysin signal sequence was included in the construction. The expressed protein is expected to have the sequence Ala-Tyr-Pro-Leu... (i.e., an extra alanine residue on the N-terminus of the proenzyme). Thirty cycles (94 °C, 30 s; 55 °C, 1 min; 72 °C, 2 min) of PCR were used to generate a 787 bp fragment encoding the described 83 amino acid prosequence and 173 amino acids of mature stromelysin. The fragment was digested with *Nco*I and *Bam*HI, and ligated into the *Nco*I and *Bam*HI sites of the T7 expression vector pET-8c (Studier et al., 1990). Initial clones containing the stromelysin sequence were obtained by transformation of the nonexpressing *E. coli* strain

Table I: Purification of Pro-sfSTR from a 5-L Shake Flask Culture of *E. coli* BL21(DE3)[pET-sfSTR]

step	protein			activity[a]			
	volume of fraction (mL)	concn (mg/mL)	total amount (mg)	total amount (units)[b]	specific activity (units/mg)	yield (%)	purification factor (*x*-fold)
crude extract	45	14.8	664	54720	82	100	1
40–50% ammonium sulfate	45	4.6	207	37710	182	69	2.2
Q-Sepharose	120	0.43	52	34600	665	63	8.1
S-Sepharose	72	0.57	41	35280	860	64	10.5

[a] Pro-sfSTR was activated with APMA for 12 h as described under Materials and Methods, and assayed with Nle-SP. [b] One unit is defined as 1 nmol of Nle-SP hydrolyzed per minute at 25 °C.

BL21. The double-stranded plasmid was sequenced by using the dideoxy method (Sanger et al., 1977) with recombinant T7 DNA polymerase (United States Biochemical, Cleveland, OH). A PCR error was detected which introduced a Phe-209 to Leu alteration in the coding sequence of stromelysin. To restore the wild-type sequence, a 587 bp *Xba*I–*Hin*dIII fragment containing the mutation was removed and replaced with the same fragment from pC4. Sequencing of the resulting plasmid pET-psfSTR showed no differences from the original cDNA clone. *E. coli* strain BL21(DE3) was transformed with pET-psfSTR to allow for the T7 RNA polymerase mediated transcription of the subcloned, short form of prostromelysin.

Pro-sfSTR Purification (Table I). (i) Crude Cell Extract. A 5-L culture of *E. coli* strain BL21(DE3)[pET-sfSTR] was grown in M9 media supplemented with 0.1% casamino acids and 100 μg/mL ampicillin at 37 °C with shaking. When a cell density corresponding to an OD_{600} of 0.8 was reached, IPTG was added to a concentration of 0.1 mM, and the incubation was continued for another 4 h at 27 °C. All further steps were performed at 4 °C. Cells were pelleted at 4400*g* and resuspended in 20 mL of 100 mM Tris-HCl (pH 8.0), 5 mM $CaCl_2$, 0.5 mM $Zn(OAc)_2$, 0.05% Brij, and 1 μg/mL each of leupeptin, aprotinin, and pepstatin. Cell extracts were routinely stored frozen at −20 °C at this point for periods up to 1 week. Frozen cells were thawed and lysed by two passes through a French press at 20 000 psi. The resulting viscous lysate was diluted with the resuspension buffer to 45 mL and treated for 30 s with a Polytron homogenizer (Brinkmann). Insoluble material was removed by centrifugation at 100000*g* for 1 h, and the resulting supernatant was used immediately for further purification.

(ii) Ammonium Sulfate Precipitation. All subsequent purification operations were performed in the cold room at 4 °C. To the 45 mL of clarified crude extract obtained in (i) was slowly added 9.68 g of ammonium sulfate (40% saturation). After the mixture was stirred for 10 min, the supernatant was collected from a 20-min centrifugation at 28000*g*. Ammonium sulfate (2.52 g) was added to the supernatant to give 50% saturation, and the protein pellet was separated by centrifugation at 70000*g* for 20 min. The pellet was resuspended to 45 mL in 20 mM Tris-HCl (pH 8.0) containing 5 mM $CaCl_2$ and 0.02% NaN_3.

(iii) Q-Sepharose Anion-Exchange Chromatography. A Pharmacia XK-26 column containing 50 mL of Q-Sepharose (9.4 × 2.6 cm) was equilibrated with 20 mM Tris-HCl (pH 8.0) containing 5 mM $CaCl_2$ and 0.02% NaN_3 [adjusted to a conductivity of 7.0 mS ($m\Omega^{-1}$/cm) with 4 N NaCl]. The feed for the column was prepared by filtering the 45 mL of resuspended pellet obtained in (ii) through a Millipore 0.45-μm filter (Syringe-type) and adjusting the conductivity to 7.0 mS by diluting with approximately 90 mL of 10 mM Tris-HCl (pH 8.0) containing 5 mM $CaCl_2$. This feed was loaded onto the column at a flow rate of 3.5 mL/min. The column was then washed (3.5 mL/min) with 250 mL of equilibration buffer, and 25-mL fractions were collected. Next, a 750-mL linear gradient of 0–0.3 M NaCl in equilibration buffer was applied to the column (0.75 mL/min), and 10-mL fractions were collected. Fractions were assayed by immunoblot analysis of a SDS–polyacrylamide gel [anti-prostromelysin antiserum provided by M. Lark (Merck)]. Fractions containing pro-sfSTR were pooled, adjusted to 20 mM MES with 1.0 M MES (pH 6.5), and then adjusted to a conductivity of 5.0 mS with 5 mM MES (pH 6.5) containing 5 mM $CaCl_2$, and finally to pH 6.5 with HCl or NaOH as required.

(iv) S-Sepharose Cation-Exchange Chromatography. A Pharmacia XK-16 column containing 40 mL of S-Sepharose (20 × 1.6 cm) was equilibrated with 20 mM MES (pH 6.5) containing 5 mM $CaCl_2$ and 0.02% NaN_3. The pH- and conductivity-adjusted pool from (iii) was loaded at a flow rate of 1.4 mL/min. The column was washed with 200 mL of equilibration buffer, and prostromelysin was eluted with a 600-mL linear gradient of 0–0.3 M NaCl in the pH 6.5 equilibration buffer at a flow rate of 0.3 mL/min. Eight-milliliter fractions were collected and assayed for pro-sfSTR by SDS–polyacrylamide gel electrophoresis. A pool of the fractions of highest purity was concentrated to 1–2 mg/mL by using an Amicon stirred cell equipped with a YM-5 ultrafiltration membrane. The purified proenzyme was stored at 4 °C for not more than 2 days before activation.

Activation of Pro-sfSTR. (i) APMA. A 20 mM stock solution of (aminophenyl)mercuric acetate (APMA) was freshly prepared as described by Stricklin et al. (1983). APMA was dissolved in 50 mL of 0.1 N NaOH and then adjusted to pH 11 with 5 N HCl. Pro-sfSTR at a concentration of 1–2 mg/mL was buffered with 0.1 M Tris-HCl (pH 7.5) and incubated with one-tenth volume of the 20 mM APMA stock at 37 °C. Aliquots were removed as a function of time, and the activation was monitored by both SDS gel electrophoresis and Nle-SP activity assay (see Kinetic Characterization). After activation was complete, APMA and propeptide fragments were separated from mature sfSTR by gel filtration chromatography on Sephacryl S-100 HR. A Pharmacia XK-16/70 column containing 120 mL of Sephacryl S-100 HR was equilibrated in 20 mM Tris-HCl (pH 7.5) containing 5.0 mM $CaCl_2$ and 0.02% NaN_3; 5.5 mL of the APMA-activated S-Sepharose pool was loaded onto the column and eluted with 140 mL of the equilibration buffer at a flow rate of 0.3 mL/min. Fractions of 2-mL volume were collected and assayed by gel electrophoresis.

(ii) Trypsin Digestion. Purified pro-sfSTR, at a concentration of 2.5 μM in 20 mM MES (pH 6.5) containing 5 mM $CaCl_2$ and 0.02% NaN_3, was incubated with 25 nM trypsin at 37 °C for 30 min. Trypsin was removed by the addition of a 4-fold molar excess of water-washed soybean trypsin inhibitor bound to agarose (Sigma, St. Louis, MO) and centrifugation of the trypsin–inhibitor complexes for 3 min at 12000*g*.

Protein Analysis. Protein concentrations were determined

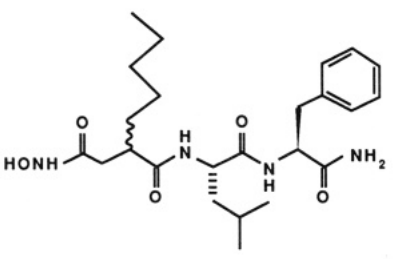
as described by Bradford (1976) using a kit from Bio-Rad. Protein samples were analyzed by using SDS–polyacrylamide gels as described by Laemmli (1972). Protein bands were detected with either Coomassie Brilliant Blue R-250 or with reagents in a Daiichi silver-staining kit from Integrated Separation Systems (Hyde Park, MA). N-Terminal amino acid sequencing was performed on proteins that were separated by gel electrophoresis and transferred onto poly(vinylidene difluoride) (PVDF) membranes (Millipore, Bedford, MA) according to the procedure described by Towbin (1979). Membranes were washed extensively with distilled water, the protein bands were stained with Coomassie Brilliant Blue, and the appropriate bands were cut out. Proteins immobilized on membranes were sequenced under gas-phase conditions according to the manufacturer's instructions with either a Porton (Tarzana, CA) 2020 or an Applied Biosystems (Foster City, CA) 477A sequencer equipped with on-line PTH amino acid separation.

Kinetic Characterization. (*i*) *Substance P Substrate Assay.* A semicontinuous, HPLC-based assay of enzymatic activity was performed as previously described (Harrison et al., 1989; Teahan et al., 1989). In a typical kinetic measurement, 1.7 mL of assay buffer consisting of 20 mM Tris-HCl (pH 7.5), 10 mM $CaCl_2$, 0.02% NaN_3, 0.05% Brij-35, and varying concentrations of Nle-SP (Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Nle-NH_2) was thermally equilibrated to 25 °C for 5 min in the temperature-controlled carousel of a Hitachi Model 655A-40 autosampler. The hydrolysis of Nle-SP was initiated by the addition of 30 μL (1.3 μg, 66 pmol) of trypsin-activated stromelysin. At 4-min intervals, 30-μL aliquots were withdrawn by the autosampler and injected onto a Whatman RAC II Partisil 5 C8 column (10 cm × 4.6 mm i.d.). Elution was isocratic (flow rate = 1.5 mL/min) with a mobile phase of 0.1% TFA in CH_3CN/H_2O (33:67, v/v), and absorbance was measured at 215 nm using a Kratos Spectroflow 757 variable-wavelength detector. The C-terminal hydrolysis product, fragment 7–11 (Nle-SP^{7-11}), is the last peptide eluted from the column (at approximately 3 min). Quantitation of the reaction rates was obtained by integration of the Nle-SP^{7-11} peak with a Nelson Turbochrome Data Station integrator. Kinetic parameters and their standard error were obtained by a nonlinear least-squares fit to the Michaelis–Menten equation.

(*ii*) *Inhibition by Hydroxamate.* Reaction rates were determined as described above (i). Nle-SP concentration was varied from 50 μM to 1 mM at different fixed levels of hydroxamate **I** (ICI-U24278; DiPasquale et al., 1986) varying from 25 nM to 2.5 μM. The K_i values for full-length stromelysin and sfSTR were determined from a nonlinear least-squares fit to the equation:

$$v = VA/[K(1 + I/K_i) + A]$$

where v = the rate of hydrolysis of Nle-SP, $V = V_{max}$, K = the K_m for Nle-SP, A = the concentration of Nle-SP, and I = the concentration of hydroxamate **I**.



I

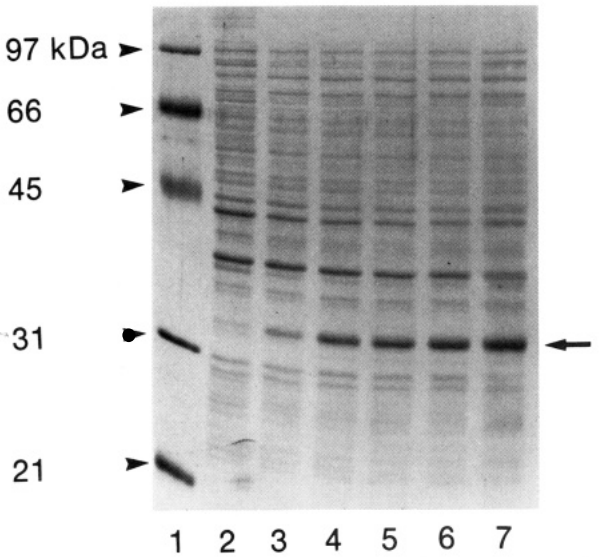


FIGURE 2: Expression time course for pro-sfSTR from *E. coli* as followed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (12%) of whole cell extracts. Lane 1, marker proteins. Lane 2, uninduced culture. Lanes 3–7, 1, 2, 3, 4, and 6 h after induction, respectively. Molecular weight standards used were phosphorylase (97 400), bovine serum albumin (66 200), ovalbumin (45 000), carbonic anhydrase (31 000), and soybean trypsin inhibitor (21 500).

Results

E. coli Expression of Pro-sfSTR. E. coli BL21(DE3) cells contain a genomic copy of the T7 RNA polymerase gene under control of the *lac* promoter. Addition of IPTG to BL21(DE3) transformed with the plasmid pET-psfSTR induces the production of T7 RNA polymerase and the expression of pro-sfSTR. The levels and time course of expression of pro-sfSTR were analyzed by SDS–PAGE of BL21(DE3)[pET-psfSTR] cell lysates (Figure 2). A novel protein migrating with an apparent molecular weight of 33K appeared 1 h after induction and increased to a maximum level within 4–6 h. Although this protein had a molecular weight by gel electrophoresis somewhat larger than the size predicted from the cDNA sequence (29K), it reacted with anti-stromelysin antiserum on Western blots (data not shown). To further characterize this short form of recombinant prostromelysin, the protein was purified from induced cultures.

Purification of Pro-sfSTR. The purification of pro-sfSTR was accomplished by a three-step protocol (Table I) and yielded 41 mg of homogeneous proenzyme from 5 L of bacterial culture. After an initial ammonium sulfate fractionation, the resuspension was applied to a Q-Sepharose column, and the proenzyme was eluted in the wash fractions. Most of the contaminating proteins were removed in this step (Figure 3, lane 5). Truncated prostromelysin bound to the subsequent S-Sepharose column and was eluted at 0.12 M NaCl at greater than 98% purity (Figure 3, lane 6). The minor contaminant at 24 kDa reacts with anti-stromelysin antiserum and corresponds to mature sfSTR.

Activation and Characterization of Pro-sfSTR. Prostromelysin can be activated by organomercurials that perturb the structure of the enzyme and allow self-cleavage. Thus, the ability of stromelysin to become activated by APMA indicates that the enzyme is properly folded and functional. The time courses of APMA activation for natural, full-length prostromelysin and pro-sfSTR were examined by SDS–PAGE and revealed similar activation rates (Figure 4). Longer APMA treatment of the full-length enzyme resulted in the appearance of lower molecular weight forms at approximately 28K. Conversely, the activated short form showed no further deg-

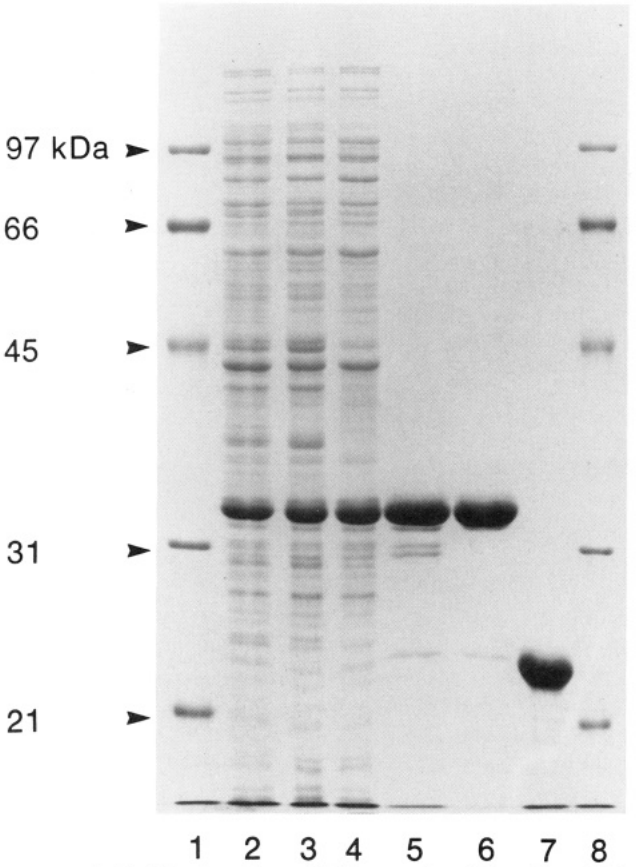


FIGURE 3: Purification of pro-sfSTR as monitored by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (12%). Lane 1, marker proteins. Lanes 2–7 each contain 10 μg of total protein. Lane 2, whole cell extract of BL21(DE3)[pET-psfSTR] 6 h after induction by IPTG. Lane 3, 0–40% ammonium sulfate supernatant of the crude extract. Lane 4, 40–50% ammonium sulfate precipitation. Lane 5, pooled pro-sfSTR wash fractions after Q-Sepharose chromatography. Lane 6, pooled pro-sfSTR fractions after S-Sepharose chromatography. Lane 7, APMA-activated pro-sfSTR after Sephacryl S-100 HR gel filtration. Lane 8, marker proteins (same as in Figure 2).

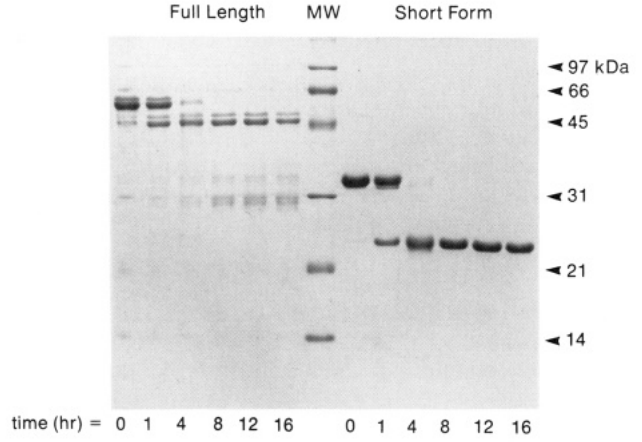


FIGURE 4: Comparison of APMA activation of full-length prostromelysin and pro-sfSTR. 1.5 mg/mL samples of full-length and short-form prostromelysin were activated with 2 mM APMA for the times indicated at 37 °C. Aliquots (5 μg) were removed and analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (15%).

radation and was stable at 24 kDa. Analogous to the natural, full-length prostromelysin, the purified proenzyme, pro-sfSTR, can be completely converted to a species lacking the 10-kDa propeptide with trypsin (30 min, not shown). The results of preliminary assays, which monitored solubilization of [^{3}H]-transferrin, indicated that the short-form enzyme had similar proteolytic activity to the full-length enzyme. For a more detailed comparison of these two forms, Nle-SP was used as a substrate. The product of the trypsin digestion of pro-sfSTR, sfSTR, showed identical cleavage specificity (data not shown) and similar kinetic parameters for Nle-SP as full-length stromelysin (Table II). In addition, sfSTR and full-length stromelysin showed similar K_i values (42 and 120 nM, respectively) for the competitive inhibitor, hydroxamate **I**.

Table II: Comparison of Substrate and Inhibitor Kinetic Parameters for Full-Length and Short-Form Stromelysin[a]

	full length	short form
k_{cat}	54 ± 9	81 ± 6
K_m	0.9 ± 0.3	1.4 ± 0.2
k_{cat}/K_m	60 ± 22	58 ± 9
K_i	120 ± 10	42 ± 6

[a] Substrate kinetic parameters k_{cat}, K_m, and k_{cat}/K_m were determined for Nle-SP and have units of min^{-1}, mM, and mM^{-1} min^{-1}, respectively. K_i inhibition values, determined for hydroxamate **I**, have units of nM.

N-Terminal Sequence Analysis. The N-terminal sequences of both the proenzyme, pro-sfSTR, and the mature enzyme after activation with APMA, sfSTR, were determined from preparations that were separated by SDS–polyacrylamide gel electrophoresis and transferred onto PVDF membranes. The N-terminal residues of the proenzyme were found to be Ala-Tyr-Pro-Leu, in agreement with the sequence predicted by the cloned insert minus the initiating methionine. When the second residue of a bacterially expressed recombinant protein is an alanine, the bacterial methionine aminopeptidase efficiently removes the N-terminal methionine (Flinta et al., 1986; Hirel et al., 1989). The N-terminal sequence of mature sfSTR was determined to be Phe-Arg-Thr-Phe. This is identical with the N-terminal sequence previously described for the full-length enzyme after APMA activation (Whitham et al., 1986; Nagase et al., 1990).

Stability. To compare the stability of the APMA-activated short form of stromelysin with that of similarly activated full-length enzyme, purified preparations of each were stored at 4, 23, and 37 °C. Aliquots were removed and analyzed by silver staining of an SDS–polyacrylamide gel. Under the activation conditions used, approximately 90% of the short-form and full-length proenzymes was activated. During the activation period, a substantial fraction of full-length stromelysin was converted to a 24–32-kDa group of proteins (Figure 5, lane 1), while no forms smaller than that predicted by the truncated primary sequence of the short-form enzyme were observed (Figure 5, lane 2). Upon storage at 23 and 37 °C, further degradation of the full-length enzyme was seen (Figure 5, lanes 5 and 7), while the short-form remains intact (Figure 5, lanes 6 and 8). Short-form enzyme showed a barely detectable degradation after 90 days at 4 °C (Figure 5, lane 10). Conversely, the full-length stromelysin was nearly completely converted to a heterogeneous mixture of lower molecular weight forms under the same conditions (Figure 5, lane 9).

Discussion

The activation of human fibroblast prostromelysin is accompanied by a decrease in size from a molecular weight of 57K to 45K as analyzed by SDS–PAGE. This is consistent with the loss of the propeptide (Figure 6). In a concentration-dependent manner, indicative of an autolytic bimolecular reaction, the 45-kDa form is further processed to a heterologous group of polypeptides centered at 28 kDa (Okada et al., 1986, 1988, 1989; Murphy et al., 1987; Nagase et al., 1990) without a concomitant loss in enzymatic activity. The presence

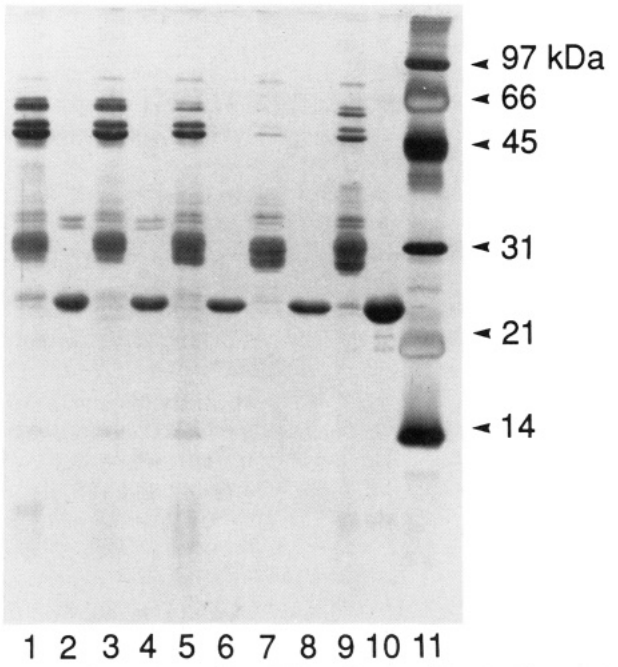


FIGURE 5: Comparison of the stability of active sfSTR and full-length stromelysin as monitored by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (12%) and visualization by silver staining. APMA-activated (37 °C, 12 h) samples at 0.5 mg/mL were incubated at the indicated temperature for the given period of time, and then frozen at −30 °C pending gel analysis. Lane 1, full length immediately after APMA activation. Lane 2, sfSTR immediately after APMA activation. Lane 3, full length (4 °C, 7 days). Lane 4, sfSTR (4 °C, 7 days). Lane 5, full length (23 °C, 7 days). Lane 6, sfSTR (23 °C, 7 days). Lane 7, full length (37 °C, 7 days). Lane 8, sfSTR (37 °C, 7 days). Lane 9, full length (4 °C, 90 days). Lane 10, sfSTR (4 °C, 90 days). Lanes 1–8 contain 3.8 μg of total protein; lanes 9 and 10 contain 5.7 μg of total protein. Lane 11, marker proteins [same as in Figure 2, plus lysozyme (14 400)].

of 28-kDa forms has not been detected in vivo, even in synovial fluids containing elevated stromelysin levels, suggesting that there may not be any biological significance associated with the self-cleavages observed with purified and concentrated samples of the enzyme (M. Lark, unpublished data). Nevertheless, the relatively high protein concentrations required for biophysical studies result in unacceptable heterogeneity in the sample. In fact, this degradation is already detected during the in vitro activation process (Figure 4, full-length, 8 h). N-Terminal sequence analysis on the mixture of lower molecular weight forms centered at 28 kDa indicated that each of these species of approximately 28 kDa has the same amino terminus as the 45-kDa form (Nagase et al., 1990). These observations have led to the suggestion that the 45-kDa form of stromelysin is folded into two domains, a N-terminal, catalytic domain and a C-terminal domain of unknown function (Sanchez-Lopez et al., 1988). The two domains are connected by a proteolytically labile hinge region which is susceptible to bimolecular, self-cleavages (Figure 7). These cleavages occur at many locations within this region and, since the C-terminal heterogeneity is not reduced with time, the hinge region appears to be a much better substrate in the bilobal structure than after a single cleavage in this region releases the C-terminal domain. The existence of these imprecise cleavages became particularly apparent and troublesome when protein concentrations were increased to the greater than 1 mg/mL levels required for biophysical studies such as X-ray crystallography and macromolecular NMR.

As a solution to this heterogeneity problem, and to directly test the two-domain model of stromelysin, a strategy of expression of a deliberately, C-terminally shortened form of the stromelysin gene was pursued. We reasoned that if a termination site was engineeered to lie to the N-terminal side of the hinge region, and if the C-terminal domain is not required for the proper folding of the catalytic domain, a suitably stable form of active stromelysin may be created. The catalytic domain must, of course, include the amino acid residues believed to be important in liganding the catalytic zinc. Two of these residues are most likely the histidines (His-201 and His-205) found within the conserved "zinc box" (Vallee & Auld, 1990) that has sequence homology with thermolysin, a metalloprotease with known three-dimensional structure (Colman et al., 1972). The location of the third ligand within the mature sequence, possibly an aspartate or a glutamate based on analogy with thermolysin, is unknown. The remaining ligand is thought to be contributed by Cys-75 of the propeptide. Three additional considerations guided us in our choice of the point of termination for the putative catalytic domain. The first was the discovery of PUMP, a metalloprotease with 55% sequence identity with stromelysin, but completely lacking the C-terminal domain (Muller et al., 1988; Quantin et al., 1989). Its deduced amino acid sequence and alignment with stromelysin are shown in Figure 6. Second, a recombinant transin (rat stromelysin) enzyme has been constructed which lacks this domain and retains protease activity (Sanchez-Lopez et al., 1988). This truncation is just after the amino acid corresponding to Leu-257 of human stromelysin. Lastly, an active N-terminal fragment of collagenase, a related metalloproteinase with greater than 50% homology to stromelysin (Figure 6), is produced autocatalytically by cleavage after the amino acid corresponding to Pro-256 of stromelysin (Whitham et al., 1986; Clark & Cawston, 1989). Unlike stromelysin, however, the cleavage occurs in a precise fashion, and the C-terminal fragment is stable toward further digestion. Taking all this information together, it was decided to construct an analogue of stromelysin exactly corresponding to the PUMP homologue, that is, eliminate the C-terminal 205 amino acids.

Though mammalian cell culture expression systems for full-length human stromelysin (Murphy et al., 1987) and full-length or C-terminally truncated forms of rat stromelysin (Sanchez-Lopez et al., 1988) had been established, the development of an *E. coli* expression system potentially offered many advantages. First, heterologous expression of recombinant proteins in *E. coli* allows expressing cultures to be generated in a short period of time, usually less than 6 h, and in a very simple growth apparatus such as an incubated shake flask. This quick and convenient generation of expressing cultures allows the timely evaluation of site-directed mutants necessary for establishing structure/activity correlations. Second, since *E. coli* are not equipped to glycosylate proteins, heterogeneity due to partial glycosylation is eliminated. In the case of stromelysin, the glycosylation occurring in mammalian cells is known to be nonessential for enzymatic activity (Okada et al., 1988). Finally, the culture conditions of an *E. coli* expression system are easily modified to allow either specific or uniform isotopic incorporation (^{13}C, ^{15}N, ^{2}H, etc.) into the expressed protein. This is a particularly important consideration for multidimensional NMR studies. Subcloning the truncated prostromelysin gene into the pET expression vector resulted in excellent overproduction of a soluble, cytoplasmic protein accounting for approximately 10% of the total cellular protein of *E. coli*. As a result, over 40 mg of the homogeneous short form of prostromelysin, pro-sfSTR, was prepared from 5 L of culture.

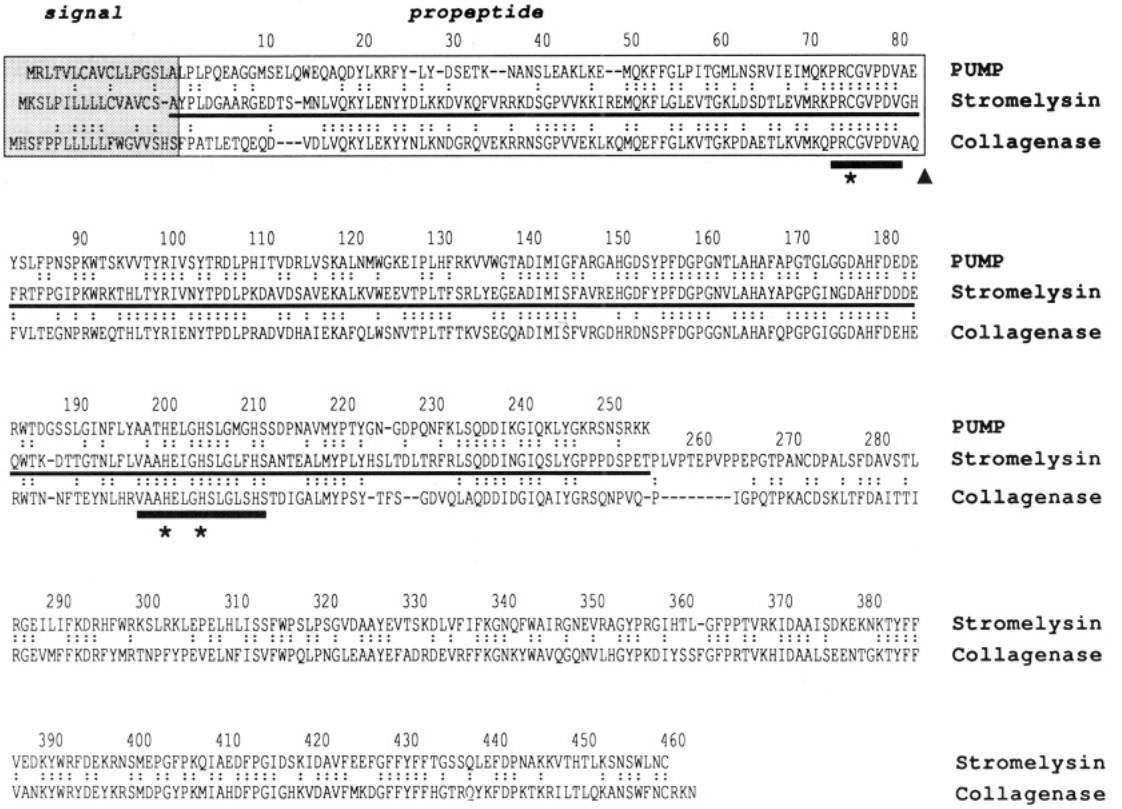


FIGURE 6: Comparison of the amino acid sequences of PUMP, stromelysin, and collagenase. The numbering of residues is for the stromelysin sequence with the first residue of the propeptide designated as number 1. Amino acids in PUMP and collagenase conserved with those in stromelysin are marked by colons. The triangle denotes the ultimate cleavage site in stromelysin following APMA activation. The residues most likely to ligand the active-site zinc are marked by asterisks and are found within highly conserved regions (bold underlined). The mammalian signal sequences are shaded, and the propeptides are boxed. The underlined stromelysin sequence corresponds to the *E. coli* expressed pro-sfSTR. PUMP and collagenase sequences are from Muller et al. (1988).

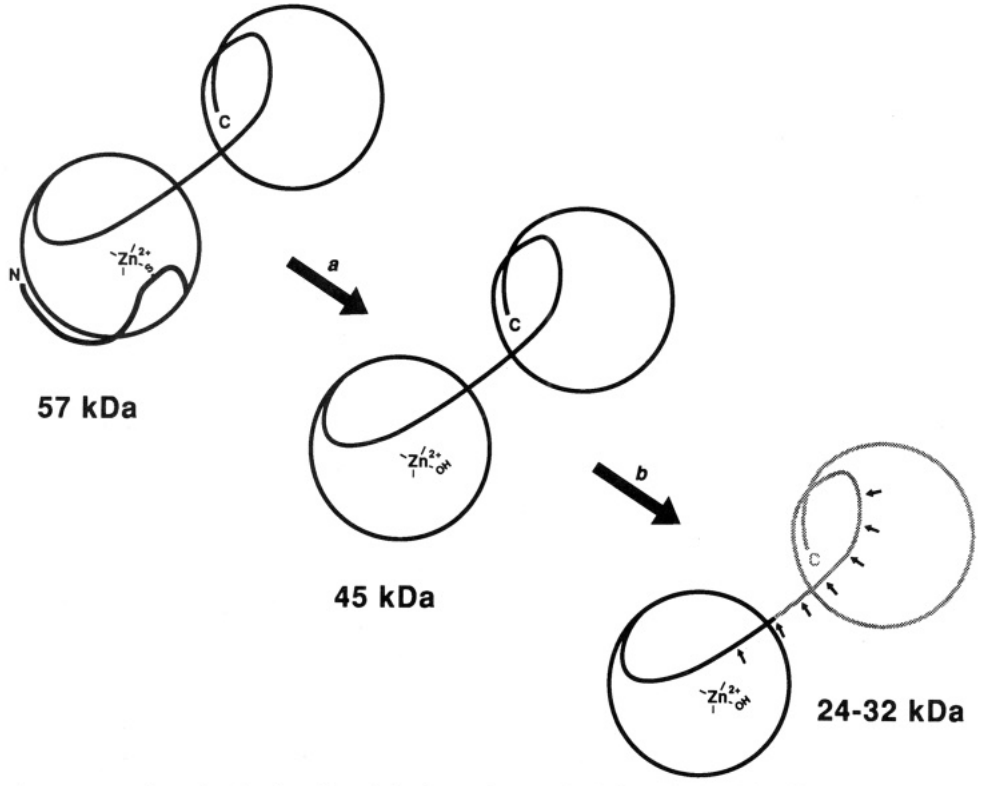


FIGURE 7: Schematic representation of activation (a) and the imprecise, autolytic degradation (b) of full-length stromelysin. The activation process (a) involves the proteolytic release of the 82 amino acid propeptide. Concentrated solutions of the 45-kDa form result in a bimolecular, self-cleavage reaction (b, at the arrows) to give a mixture of active forms ranging in molecular weight from 24K to 32K.

Initial indications that the pro-sfSTR was properly folded came from the results of APMA treatment. Studies on the activation of full-length prostromelysin have suggested that APMA complexing with Cys-75 of the propeptide causes a

conformation change which leads to an intramolecular cleavage at the Glu[68]–Val[69] bond of the propeptide, followed by an intermolecular cleavage at the His[82]–Phe[83] bond (Nagase et al., 1990). Obviously, only correctly folded material should display such activation and propeptide removal. The relative reaction rates for this activation with full-length prostromelysin and pro-sfSTR were similar (Figure 4), supporting the concept of a bilobal model (Figure 7). Further confirmation of the structural similarity of the active sites of full-length and sfSTR is provided by the examination of the kinetic parameters for hydrolysis of substrate and inhibition by a peptide-like hydroxamate. Within experimental error, full-length and sfSTR behaved identically in these assays, displaying the same k_{cat} and K_m values for hydrolysis of Nle-SP and nearly identical K_i values for compound I (Table II). The C-terminal domain may contribute specific interactions with macromolecular substrates such as collagen, proteoglycan, or fibronectin. This possibility is currently being investigated by comparing the reactivity of sfSTR and full-length stromelysin toward these macromolecular substrates.

Since the C-terminally truncated form of stromelysin, sfSTR, provided similar binding interactions to peptide substrates and small-molecule inhibitors as natural, full-length stromelysin, the increased homogeneity afforded by elimination of the proteolytically susceptible C-terminal domain and hinge region should allow a structural analysis leading to a relevant three-dimensional picture for inhibitor design. In particular, a three-dimensional structure of a complex of sfSTR with an inhibitor, such as I, may suggest modifications to improve inhibitor potency, and/or suggest positions of the inhibitor tolerant of modifications to improve pharmacokinetics.

ACKNOWLEDGMENTS

We thank Dr. M. Lark for generously providing anti-stromelysin antibody, N. Santoro and E. Butkiewicz for expert technical assistance, and Drs. M. Lark, R. Stein, J. Schmidt, and J. Springer for encouragement and helpful discussions.

REFERENCES

Bradford, M. M. (1976) *Anal. Biochem. 72*, 143–147.
Brinckerhoff, C. E., Ko, S., Mitchell, T. I., Oram, F., Coon, C. I., Palmiter, R. D., & Nagase, H. (1990) *J. Biol. Chem. 265*, 22262–22269.
Chin, J. R., Murphy, G., & Werb, Z. (1985) *J. Biol. Chem. 260*, 12367–12376.
Clark, I. M., & Cawston, T. E. (1989) *Biochem. J. 263*, 201–206.
Colman, P. M., Jansonius, J. N., & Matthews, B. W. (1972) *J. Mol. Biol. 70*, 701–724.
DiPasquale, G., Caccese, R., Pasternak, R., Conaty, J., Hubbs, S., & Perry, K. (1986) *Proc. Soc. Exp. Biol. Med. 183*, 262–267.
Flinta, C., Persson, B., Jornvall, H., & von Heijne, G. (1986) *Eur. J. Biochem. 154*, 193–196.
Galloway, W. A., Murphy, G., Sandy, J. D., Gavrilovic, J., Cawston, T. E., & Reynolds, J. J. (1983) *Biochem. J. 209*, 741–752.
Gowen, M., Wood, D. D., Ihrie, E. J., Meats, J. E., & Russell, R. G. G. (1984) *Biochim. Biophys. Acta 797*, 186–193.
Harrison, R., Teahan, J., & Stein, R. (1989) *Anal. Biochem. 180*, 110–113.
He, C., Wilhelm, S. M., Kronberger, A., Bauer, E. A., Grant, G. A., & Eisen, A. Z. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2632–2636.
Hirel, P.-H., Schmitter, J.-M., Dessen, P., Fayat, G., & Blanquet, S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8247–8251.
Ito, A., & Nagase, H. (1988) *Arch. Biochem. Biophys. 267*, 211–216.
Johnson, B. A., McClain, S. G., Doran, E. R., Tice, G., & Kirsch, M. A. (1990) *BioTechniques 8*, 424–429.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
MacNaul, K. C., Chartrain, N., Lark, M., Tocci, M. J., & Hutchinson, N. I. (1990) *J. Biol. Chem. 265*, 17238–17245.
Matrisian, L. M. (1990) *Trends Genet. 6*, 121–125.
Muller, D., Quantin, B., Gesnel, M.-C., Millon-Collard, R., Abescassis, J., & Breathnach, R. (1988) *Biochem. J. 253*, 187–192.
Murphy, G., McAlpine, C. G., Poll, C. T., & Reynolds, J. J. (1985) *Biochim. Biophys. Acta 831*, 49–58.
Murphy, G., Cockett, M. I., Stephens, P. E., Smith, B. J., & Docherty, A. J. P. (1987) *Biochem. J. 248*, 265–268.
Nagase, H., Enghild, J. J., Suzuki, K., & Salvesen, G. (1990) *Biochemistry 29*, 5783–5789.
Okada, Y., & Nakanishi, I. (1989) *FEBS Lett. 249*, 353–356.
Okada, Y., Nagase, H., & Harris, E. D. (1986) *J. Biol. Chem. 261*, 14245–14255.
Okada, Y., Harris, E. D., & Nagase, H. (1988) *Biochem. J. 254*, 731–741.
Quantin, B., Murphy, G., & Breathnach, R. (1989) *Biochemistry 28*, 5327–5334.
Sanchez-Lopez, R., Nicholson, R., Gesnel, M.-C., Matrisian, L. M., & Breathnach, R. (1988) *J. Biol. Chem. 263*, 11892–11899.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Saus, J., Quinones, S., Otani, Y., Nagase, H., Harris, E. D., & Kurkinen, M. (1988) *J. Biol. Chem. 263*, 6742–6745.
Schmid, T. M., Mayne, R., Jeffery, J. J., & Linsenmayer, T. F. (1986) *J. Biol. Chem. 261*, 4184–4189.
Seltzer, J. L., Eisen, A. Z., Bauer, E. A., Morris, N. P., Glanville, R. W., & Burgeson, R. E. (1989) *J. Biol. Chem. 264*, 3822–3826.
Stricklin, G. P., Jeffery, J. J., Roswit, W. T., & Eisen, A. Z. (1983) *Biochemistry 22*, 61–68.
Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60–89.
Teahan, J., Harrison, R., Izquierdo, M., & Stein, R. (1989) *Biochemistry 28*, 8497–8501.
Towbin, H., Staehelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350–4354.
Vallee, B. L., & Auld, D. S. (1990) *Biochemistry 29*, 5647–5659.
Van Wart, H. E., & Birkedal-Hansen, H. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5578–5582.
Welgus, H. G., Jeffery, J. J., & Eisen, A. Z. (1981) *J. Biol. Chem. 256*, 9511–9515.
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. (1986) *Biochem. J. 240*, 913–916.
Wilhelm, S. M., Collier, I. E., Kronberger, A., Eisen, A. Z., Marmer, B. L., Grant, B. A., Bauer, E. A., & Goldberg, G. I. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6725–6729.